UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1997

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 7A

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

20010321 077

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1997

AQM01-06-1283

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructic ... ng and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, includi ... for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budge

AFRL-SR-BL-TR-00-

0761

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December, 1997	3. F

4. TITLE AND SUBTITLE 1997 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 7A, Armstrong Laboratory	5. FUNDING NUMBERS F49620-93-C-0063
6. AUTHOR(S) Gary Moore	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Armstrong Laboratory.

14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

GSRP FINAL REPORT TABLE OF CONTENTS i-x

APPENDICIES:

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the

1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435
1997	148	98	140	427

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1997 are shown in this table.

1997 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	490	188	32
(HBCU/MI)	(0)	(0)	(0)
GSRP	202	98	9
(HBCU/MI)	(0)	(0)	(0)
HSAP	433	140	14
TOTAL	1125	426	55

4. SITE VISITS

During June and July of 1997, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1997 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1997 SRP selected participants are shown here.

1997 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	89	123
USAF Laboratory Funds	48	9	17
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	0	0	N/A
TOTAL	9	98	140

SFRP - 188 were selected, but thirty two canceled too late to be replaced.
GSRP - 98 were selected, but nine canceled too late to be replaced.
HSAP - 140 were selected, but fourteen canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1997 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996	1997
Faculty Members	$690	$718	$740	$740	$740	$770	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % (194 out of 286) claimed travel reimbursements at an average round-trip cost of $776.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of $388. By contrast, in 1993, 58 % of SFRP associates took

orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the second year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, only 7% of the associates' information needed clarification in order for direct deposit to properly function as opposed to 10% from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1997 REPORT

The complete set of reports for the 1997 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1997 associates, as indicated below:

1997 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A -- PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	189
GSRP	97
HSAP	140
TOTAL	426

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	0	184
Master's	2	41	43
Bachelor's	0	56	56
TOTAL	186	97	298

SFRP Academic Titles	
Assistant Professor	64
Associate Professor	70
Professor	40
Instructor	0
Chairman	1
Visiting Professor	1
Visiting Assoc. Prof.	1
Research Associate	9
TOTAL	186

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B -- SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1997 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	**Evals Recv'd**	**0**	**1**	**2**	**3+**	**0**	**1**	**2**	**3+**	**0**	**1**	**2**	**3+**
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
USAFA	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
   a. Time available for evaluation of applications:
   b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5 a. Benefits of associate's work to laboratory:
   b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
   b. Enhancement of research qualifications for SFRP associate:
   c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
   b. Enhancement of knowledge for SFRP associate:
   c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*USAFA*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1997 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1997 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1997 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1997 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

I DENTIFICATION AND QUANTITATION OF N-METHYL-1-(3,4 METHYLENEDIOXYPHENYL)-2-BUTANAMINE 9MBDB) TOGETHER WITH ITS METABOLITE, 3,4-METHYLENEDIOXYPHENY.-2-BUTANAMINE (BDB) IN URINE

Benedict N. Arrey

University of Texas, San Antonio
6900 North Loop 1604 West
San Antonio, Texas 78249-0616

Final Report for
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

Armstrong Laboratory

August 1997

IDENTIFICATION AND QUANTITATION OF N-METHYL-1-(3,4-METHYLENEDIOXYPHENYL)-2- BUTANAMINE (MBDB) TOGETHER WITH ITS METABOLITE, 3,4-METHYLENEDIOXYPHENYL-2-BUTANAMINE (BDB) IN URINE.

BENEDICT N. ARREY
GRADUATE STUDENT
DEPARTMENT OF LIFE SCIENCES
UNIVERSITY OF TEXAS, SAN ANTONIO.

ABSTRACT

N-methyl -1- (3,4 -methylenedioxyphenyl) 1-2-butanamine (MBDB) is the α-ethyl homologue of 3,4 -methylenedioxymethamphetamine(MDMA), commonly known as "Ecstasy." This work accomplishes the concurrent identification and quantitation of MBDB and its more potent metabolite, 1-(3,4-methylenedioxyphenyl) -2- butanamine(BDB) using gas chromatography- mass spectrometry (GC/MS). Solid phase extraction was found to be preferable to the popular liquid-liquid extraction method. The use of 1% HCl in methanol, which is normally added after extraction to decrease the volatility of amphetamine bases, was shown to increase the overall yield of MBDB and BDB. Derivatization of **MBDB** and **BDB** with heptafluorobutyric anhydride (HFBA) permitted identification and quantitation by GC/MS in the selected ion monitoring (SIM) mode. Fragments used were m/z 176, 210, and 268 for MBDB and m/z 135, 176, and 254 for BDB. Using d_5- MDMA as an internal standard, the linear range of the method was demonstrated between 10 and 2000 ng/mL.

IDENTIFICATION AND QUANTITATION OF N-METHYL-1-(3,4-METHYLENEDIOXYPHENYL) -2-BUTANAMINE (MBDB) TOGETHER WITH ITS METABOLITE, 1- (3,4 METHYLENEDIOXYPHENYL) -2-BUTANAMINE (BDB) IN URINE.

Benedict N. Arrey

Introduction

N - methyl-1-(3,4 methylenedioxyphenyl)-2-butanamine (MBDB) is the α-ethyl homologue of 3,4 methylenedioxymethamphetamine (MDMA, "ecstasy") placed on Schedule I of Public Law 91-513 by the U.S Drug Enforcement Administration in 1985 (1). It is an entactogen which is derived from the Latin and Greek roots that means " to produce a touching within" (2). The identification and quantitation of a drug requires the proper extraction procedure which yields maximum recovery , the use of a non-interfering internal standard and finally having a reasonable number of targettable, unique molecular fragments (3). The Substance Abuse and Mental Health Services Administration (SAMHSA); formerly, the National Institute on Drug Abuse (NIDA) has stringent guidelines for laboratories to follow for accurate and precise identification of five classes of drugs of abuse. Amphetamines, cannabinoid, cocaine, opiates and phencyclidines with cut-off values of 500, 150, 15, 300, 25 ng/mL, respectively (4). However, with the continuous production of clandestine analogues, there is a need to go beyond this point to identify and quantitate other psychoactive agents, such as MBDB. Reports of urine specimens of drug users containing amounts of 0.1 - 24 μg/mL exist, but the lethal doses have not been

established. MBDB was identified and quantitated separately from other drugs of Abuse found in the specimens but 3,4- methylenedioxyphenyl-2-butanamine (BDB) was not quantitated (2).

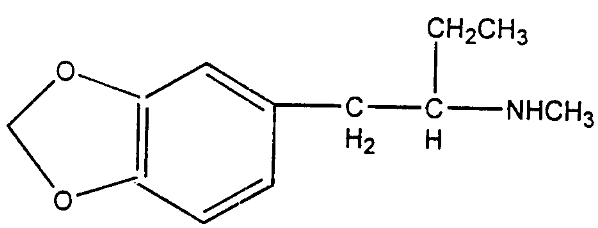
MBDB and 3,4-methylenedioxyphenyl-2-butanamine (BDB) are currently drugs claimed to have the ability to facilitate interpersonal communication by reducing the anxiety or fear that normally accompanies the discussion of emotionally painful events (5). Their presence in Europe (reported in Sweden in 1996 (2)) and in the United States indicates for development of a procedure to identify and quantitate these drugs. It is not surprising that both drugs have been found concurrently in specimens obtained from users and at autopsy, since BDB can be considered a metabolic byproduct of MBDB. It is therefore important to successfully identify and quantitate these drugs simultaneously. MBDB is one of the N-substituted analogues of 3,4-methylenedioxyamphetamine (MDA), a popular drug that has been known for years (3).

An attempt was made to use the liquid-liquid method of extraction(6), but it yielded only 50-60 % recovery. GC/MS results were inconsistent (Table 1). Linearity could not be established consistently above 1000 ng/mL. Finally, the results were not reproducible within ±20 % of known concentrations. Several internal standards were tried including d_5-MDMA, d_6-MDEA, d_5-phenylbutanamine but most of them failed to

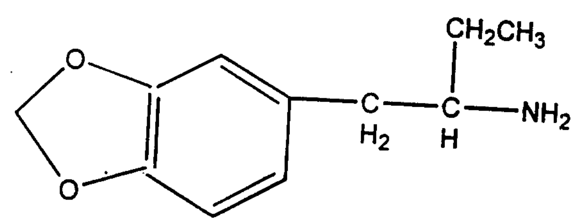
compensate for errors that occur in the sample preparation process and varying results due to GC/MS conditions.

This work addresses the extraction procedure, derivatization, use of an appropriate internal standard (d_5-MDMA) and the ability to quantitate MBDB together with its demethylated metabolite, BDB, since BDB has been known to be more potent than MBDB (7). The limits of detection (LOD), limits of quantitation (LOQ) and limits of linearity (LOL) were examined (8).

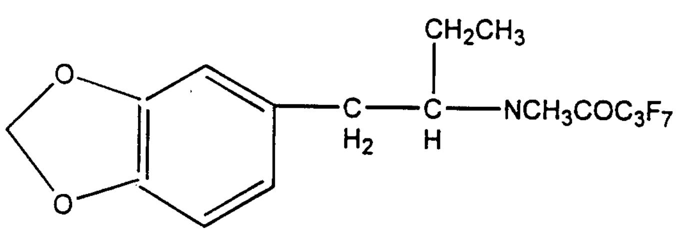
Structures



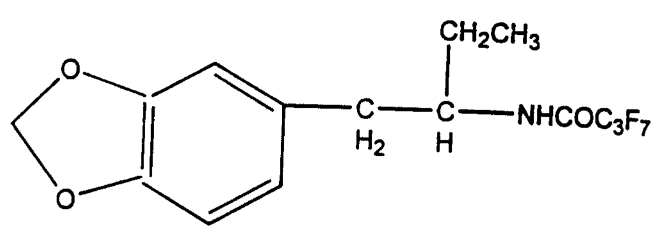
(a) MBDB underivatized



(b) BDB underivatized.



(C) MBDB derivatized



(d) BDB derivatized

Acknowledgment.

This work was accomplished with the sole sponsorship of the Air Force Office of Scientific Research in collaboration with the Drug Testing Division of Armstrong Laboratory (Forensic Investigations and Method development) Brooks Air Force Base, San Antonio; through the Graduate Student Research Program. They provided all the necessary reagents, instruments, material and technical support.

EXPERIMENTAL

Materials

The following were purchased from Research Biochemicals International (Natick, MA, USA): N-methyl-1-(3,4-methylenedioxyphenyl)-2-butanamine (MBDB) hydrochloride, 3,4-methylenedioxyphenyl-2-butanamine, 1-phenyl-2-butanamine. 1-methylenedioxiphenyl-2- methyl-d_3-aminopropane-1,2-d_2 (d_5-MDMA), 1-methylenedioxyphenyl-2-ethyl-(2, 2, 2-d_3)-aminopropane-3, 3, 3-d_3 (d_6-MDEA) were obtained from the Radian international LLC (Austin, TX, USA). Certified negative urine was obtained from UTAK Laboratories Inc. Valencia, California. Chlorobutane and Ammonium hydroxide were purchased from Aldrich Chemical Co, Milwaukee,WI.

Methanol was from Burdick and Jackson, Muskegon, MI. Heptafluorobutyric anhydride (HFBA) was from Regis Technologies, Inc. Morton Grove, IL. The following were bought from Malinckrodt chemical, Paris, Kentucky: acetic acid, sulfuric Acid, sodium hydroxide, methanol, potassium phosphate. Hydrochloric acid and ethyl acetate were from J.T.Baker, Phillipsburg, NJ. CLEAN SCREEN® extraction column cartridges (200mg/10mL), World Monitoring Extraction System, Horsham, PA.

Sample preparation and extraction

The drugs were extracted using the liquid - liquid extraction method certified by the Drug Testing Division of the Armstrong Laboratories for the extraction of amphetamines. Two milliliter aliquots of urine (samples, standards and controls) were pipetted into 50 mL screw - capped tubes. Internal standard (100 µL of d_5- MDMA) was added to each. Then, 0.3 mL 1N NaOH was added into each tube, and each was vortexed. The pH was checked to checked to insure that it was greater than 10. Ten milliliters of chlorobutane was then added to each of the samples. The tubes were capped securely and were shaken on an horizontal shaker for 10 minutes. The samples were centrifuged at 1,500 RPM for three minutes. The top organic layer of each was then transferred into new screw - capped tubes. Two milliters of 0.3 N sulfuric acid was added, and the tubes were shaken on an horizontal shaker for 10 minutes. This was followed by centrifugation at 1,500 RPM for three minutes. The top organic layer was aspirated to waste. Then,

1.0 mL 1 N NaOH was added to the bottom aqueous layer, the tubes were vortexed and again the pH was checked to be greater than 10. Ten milliliters of 1- chlorobutane was added into the tubes and they were capped and shaken again on the horizontal shaker for 10 minutes, followed by centrifugation for three minutes at 1500 RPM. The top organic layers were then transferred to clean, dry tubes. The extracts were treated with 200 µl 1% HCl in methanol. They were evaporated under a gentle stream of nitrogen in a 50 - 60°C water bath. Derivatization was done by adding 100 µl ethyl acetate and 25 µl HFBA to each extract and the mixtures were capped, vortexed and incubated in a 60-70°C water bath for 15 minutes and then evaporated to dryness under a stream of nitrogen in a 50-60°C water bath. The extracts were reconstituted with 50 µl ethyl acetate for the GC/MS analysis. Relevant gas chromatographic and mass spectrometric data are shown in Figs. 1-4.

Solid phase extraction (SPE)

The solid phase extraction procedure was also based on the method used by the Drug Testing Division of the Armstrong Laboratory. The drugs were extracted as follows: Two milliliters aliquots of urine (sample, standards and controls) were pipetted into 15 ml screw- capped centrifuge tubes. Two millimeters of 0.1M potassium phosphate buffer (pH 6.0) was added to each tube. The tubes were vortexed for twenty seconds and the pH was checked to be within 6.0±0.5. The solid phase extraction (SPE) cartridges were activated by sequentially pouring and aspirating the following solvent:

IDENTIFICATION AND QUANTITATION OF N-METHYL-1-(3,4-METHYLENEDIOXYPHENYL) -2-BUTANAMINE (MBDB) TOGETHER WITH ITS METABOLITE, 1- (3,4 METHYLENEDIOXYPHENYL) -2-BUTANAMINE (BDB) IN URINE.

Benedict N. Arrey

Introduction

N - methyl-1-(3,4 methylenedioxyphenyl)-2-butanamine (MBDB) is the α-ethyl homologue of 3,4 methylenedioxymethamphetamine (MDMA, "ecstasy") placed on Schedule I of Public Law 91-513 by the U.S Drug Enforcement Administration in 1985 (1). It is an entactogen which is derived from the Latin and Greek roots that means " to produce a touching within" (2). The identification and quantitation of a drug requires the proper extraction procedure which yields maximum recovery , the use of a non-interfering internal standard and finally having a reasonable number of targettable, unique molecular fragments (3). The Substance Abuse and Mental Health Services Administration (SAMHSA); formerly, the National Institute on Drug Abuse (NIDA) has stringent guidelines for laboratories to follow for accurate and precise identification of five classes of drugs of abuse. Amphetamines, cannabinoid, cocaine, opiates and phencyclidines with cut-off values of 500, 150, 15, 300, 25 ng/mL, respectively (4). However, with the continuous production of clandestine analogues, there is a need to go beyond this point to identify and quantitate other psychoactive agents, such as MBDB. Reports of urine specimens of drug users containing amounts of 0.1 - 24 μg/mL exist, but the lethal doses have not been

Gas chromatography/ Mass spectrometry

A model 5890 series II Hewlett - Packard gas chromatography with HP 5972 mass selective detector was used for the analysis.

A) Operating Conditions:

a)	column	HP-1
b)	Injector:	270°C
c)	Interface	280°C
d)	oven temperature	

1) Equilibration time 0.5 min

2) program 1

a)	Initial temperature	100°C
b)	Initial time	2.00 min
c)	Program rate	20/min
d)	Final temperature	200°C
e)	Final Time	0.00 min

3) Program 2

a)	Program rate	40.00 °/min
b)	Final temperature	290°C
c)	Final time	0.25

B)	Helium flow	1.0 mL/min/120°C
C)	Purge Off Time	0.00 min
D)	MSD On	3.00 min
E)	MSD Off	8.00 min
F)	Run Time	10.00 min
G)	Carrier gas	helium, 99.9995% purity

MSD PROGRAM.

DRUG	ION MONITORED	RETENTION TIMES (MIN)
BDB	135, 176, 254	6.02
MBDB	176, 210, 268	6.52
d_5-MDMA	164, 213, 258	6.18

RESULTS AND DISCUSSION.

The liquid-liquid extraction procedure for the identification and quantitation of MBDB and BDB does not produce reliable results, nor is it linear beyond 1000 ng/mL (Table 1). Moreover, there was so much inconsistency, such that solid phase extraction was tried. Other reasons for using the solid phase extraction were: (a) short extraction time required, (b) decrease use of organic solvents, and (c) the ease of the process. The solid phase extraction procedure appeared to be more than 90 % efficient. Method linearity was established between 10 ng/mL and 2000 ng/mL (Table 2 and 3). This is the range over which the analyte detected accurately reflects the quantity actually present in the aliquots, to within ± 20 %. Though the concentrations of 3000 ng/mL and 4000 ng/mL appear to be within ± 20 % of analyte concentrations, there was much data where these values were not reproducible (data not published). The overall method is still important because it allows quantitation of BDB and MBDB in the ranges of the cut-off values established by SAMHSA and the Department of Defense for amphetamines, their analogues (500 ng/mL). Currently, the lethal doses have not been established, since these drugs are always found in combination with other psychoactive drugs.

Reproducibility was established using an average of ten samples at 500 ng/mL duplicate and ten at 200 ng/mL (Table 4). The samples were extracted from two milliliters of certified negative urine using the solid phase extraction procedure as previously described. The internal standard (500 ng/mL) was added to all of the samples at the end of extraction. The extracts were dried down after addition of 200 µl of 1% hydrochloric acid in methanol (as previously described); followed by derivatization with HFBA. The percentage recovery was determined by

comparing the observed concentrations to the known standard concentrations of 500 ng/mL or 200 ng/mL and they were over 90 % in both cases. This indicates that the whole analytical procedure is reproducible using d_5 -MDMA as the internal standard. Its use as an internal standard prevents the problems one normally runs into by using the corresponding isotopes as internal standard. These problems often include: (a) interference from the isotopic hydrogen, (b) inadequate separation of internal standard from the analyte or its metabolite and (c) production of ions due to cross contamination from analyte or its metabolite.

The limits of detection (LOD) were established at 10 ng/mL. This is the point below which the analytes are no longer identifiable, based on the requirement that all ion abundance ratios must be acceptable (within ± 20 % of a known standard). The LOD was determined by analyzing decreasing concentrations of known analytes at concentrations between 40 ng/mL to 3ng/mL (Table 3). Concentrations below 9 ng/mL give ion abundance ratios which are greater than ± 20 % of those given by known standards. The limit of quantitation (LOQ) was also determined to be 10 ng/mL. This is the minimum concentration at which aliquots could be quantitated consistently within ± 20 % and the ion abundance ratios were still acceptable. It should be noted that this value is far below the cutoff values for MBDB and BDB. This assures the ability of the procedure to identify and quantitate BDB and MBDB to concentrations fifty times less than their cutoff values.

At 500 ng/mL, there is strong relative abundance of all three drugs as seen in Figs. 1, 2 and 3. This supports the fact that the following criteria for good selective ion monitoring (SIM) and derivatization were met. These include, but not limited to the ease of extraction for MBDB and BDB for analyses, a derivatization process that did not degrade the analytes, the

completeness of the derivatization with HFBA (even at low concentrations of 10 ng/mL), good chromatographic separation of MBDB, BDB and d_5- MDMA (internal standard), and "reasonable" and retention times (Fig 4). The last two criteria might be used by smaller laboratories, where retention times could be used as the basis for identifications of compounds. The fragments used for selected ion monitoring (SIM) reflect different portions of each compound. (MBDB: 176, 210, 268; BDB: 135, 176, 254; d_5-MDMA: 164, 213, 258, Fig 2,3,4). The preference of SIM over the full scan is that, in SIM, some specificity is lost but the sensitivity of the method is increased.

Extracts of MBDB and BDB (phenylethylamines) do require treatment with 1% HCl in methanol prior to evaporation of solvent . This is frequently done to decrease the volatility of amphetamines and their analogues. The avoidance of this step was found to lower recovery (Table 5). MBDB does not appear to be highly affected by nontreatment, but its demethylated metabolite BDB, shows a higher percentage variation when not treated. For concurrent identification of both analytes, it is important to include this step at all times.

CONCLUSION: This method for concurrent identification of MBDB and its metabolite, BDB, meets all the requirements for confirmation of drugs in urine, as used by many forensic and investigation laboratories. These include the reproducibility of the method, linearity, recovery and sensitivity.

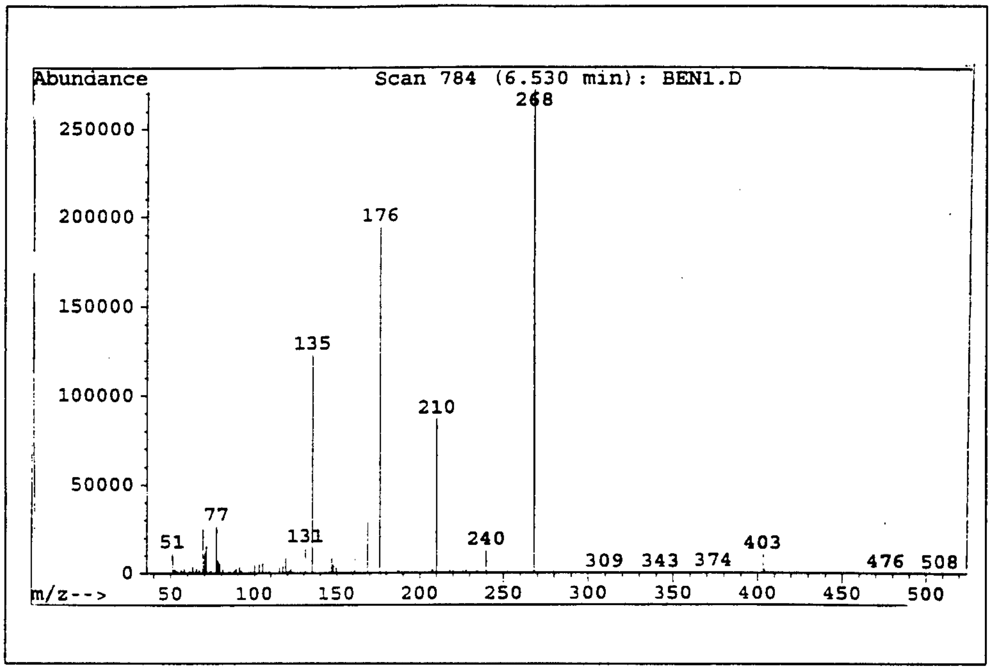


*Fig 1. **MBDB**, Relative ion abundances by selective ion monitoring(SIM)*

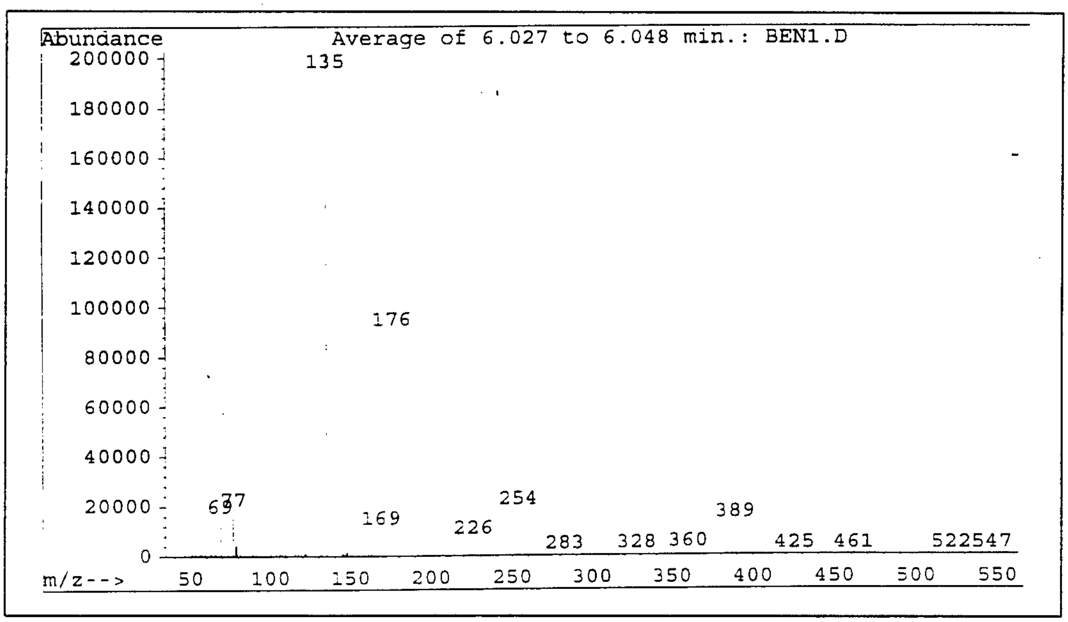


*Fig 2. **BDB**, Relative ion abundances by selective ion monitoring(SIM)*

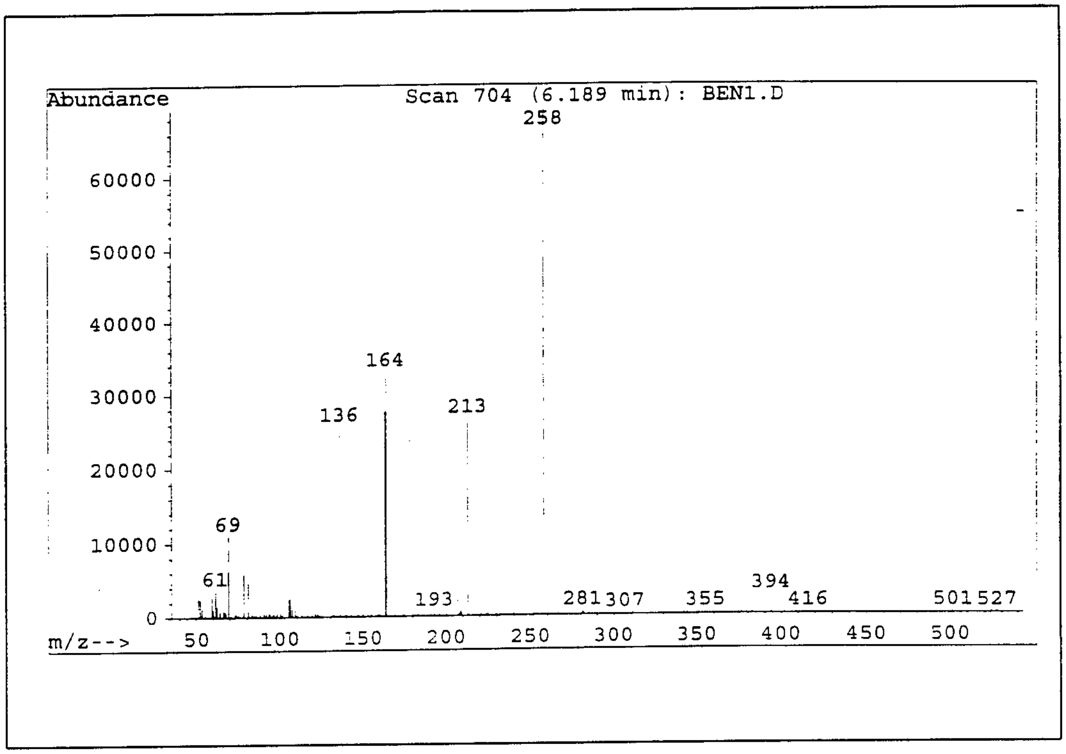


Fig 3. d_5 ***-MDMA****, Relative ion abundances by selective ion monitoring (SIM)*

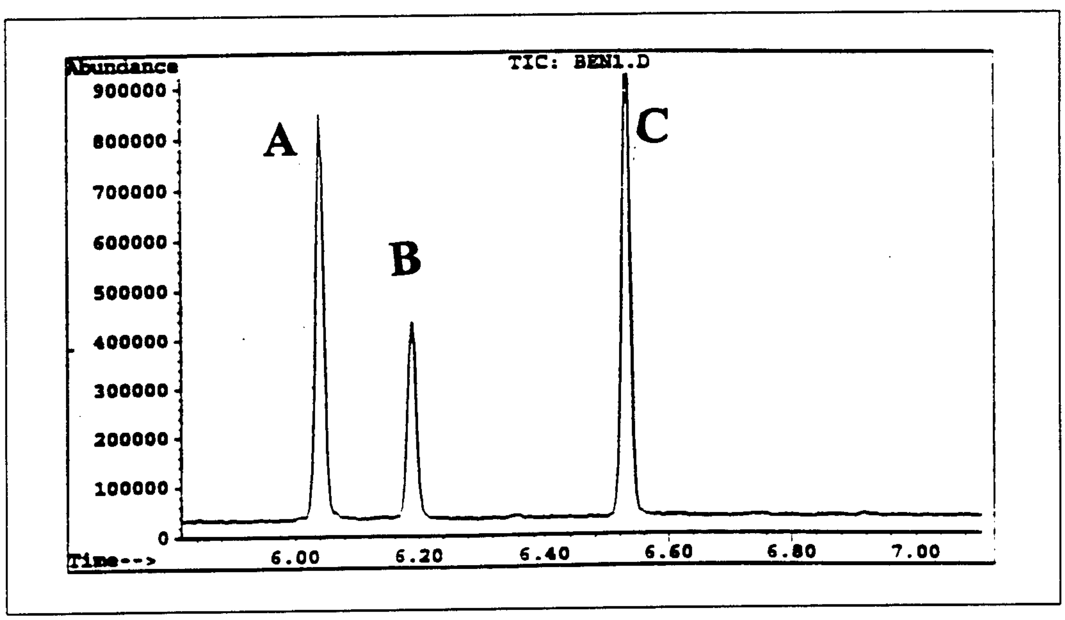


Fig 4. GC/MS chromatogram showing derivatized samples (500ng/ml). Peak identification: A, ***BDB*** *B, d_5* ***-MDMA*** *C,* ***MBDB***

TABLE 1. LINEARITY FOR MBDB/BDB BY LIQUID - LIQUID EXTRACTION.

BDB			MBDB		
THEOR CONC. ng/mL	OBSERV. CONC. ng/mL	DEV(%)	THEOR CONC. ng/mL	OBSERV. CONC. ng/mL	DEV(%)
0	0	0	0	0	0
50	55	10	50	43	-14
100	101	1	100	90	-10
200	222	11	200	173	-14
500	561	12	500	497	-1
1000	1334	33	1000	1031	3
2000	2855	43	2000	2209	10
3000	4664	55	3000	3524	17

TABLE 2.LINEARITY FOR MBDB/BDB BY SOLID PHASE EXTRACTION.

BDB			MBDB		
THEOR CONC. ng/mL	OBSERV. CONC. ng/mL	DEV(%)	THEOR CONC. ng/mL	OBSERV. CONC. ng/mL	DEV(%)
0	0	0	0	0	0
50	52.2	4	50	51	2
100	95	-5	100	94	-6
200	186	-7	200	178	-11
500	491	-2	500	467	-7
1000	1084	8	1000	1050	5
2000	2408	20	2000	2296	15
3000	3435	15	3000	3399	13
4000	4450	11	4000	4351	9

TABLE 3. LIMITS OF DETECTION/ LIMITS OF QUANTITATION BY SOLID PHASE EXTRACTION.

THEOR. CONC ng/mL	AV.BDB CONC. ng/mL	ion ratio acceptable	%DEV BDB		AV.MBDB CONC. ng/mL	ion ratio acceptable	%DEV MBDB
3	0	No	**0**		0	No	**0**
6	4.22	No	**-30**		4.38	No	**-27**
9	6.98	No	**-22**		6.75	No	**-25**
10	8.4	Yes	**-16**		8.5	Yes	**-15**
20	17.4	Yes	**-13**		17	Yes	**-15**
30	25.6	Yes	**-15**		24.3	Yes	**-19**
40	32	Yes	**-20**		32.8	Yes	**-18**

TABLE 4. REPRODUCIBILITY FOR MBDB/BDB BY SOLID PHASE EXTRACTION.

	REPRODUCIBILITY (200ng/mL)						REPRODUCIBILITY (500ng/mL)	
NO	AVER. BDB	%RECOV BDB	AVER. MBDB	RECOV. MBDB	AVER BDB	%RECOV BDB	AVER. MBDB	RECOV. MBDB
1	189	95	217	109	524	105	470	94
2	170	85	200	100	540	108	504	101
3	182	91	214	107	531	106	490	98
4	176	88	200	100	512	102	486	97
5	171	85	209	105	549	110	473	95
6	167	84	195	98	446	89	481	96
7	172	86	194	97	417	83	465	93
8	156	78	190	95	429	86	460	92
9	168	84	192	96	377	75	486	97
10	164	82	193	96	442	88	481	96
MEAN	171.5	**85.8**	200.4	**100.3**	476.7	**95.2**	479.6	**95.9**
STDEV	9.2		9.7		61.1		13.0	
%CV	5.4		4.8		12.8		2.7	

TABLE 5 . TREATMENT WITH 2% HYDROCHLORIC ACID IN MET		
THE0R	BDB	MBDB
CONC.	CONC.	CONC.
ng/mL	ng/ml	ng/ml
500	524	470
500	540	504
500	531	490
500	512	487
500	549	473
MEAN	531	485
STDEV	14	14
% CV	3	3
NO TREATMENT WITH 2 % HYDROCHLORIC ACID.		
THE0R	BDB	MBDB
CONC.	CONC	CONC.
ng/mL	ng/mL	ng/mL
500	446	481
500	417	465
500	429	461
500	377	486
500	442	481
MEAN	422	475
STDEV	30	11
% CV	7	2

References

1. T.P. Rohrig, R. Prouty. Tissue distribution of methylenedioxymethamphetamine. J. *Anal.Toxicol. **16**:52-53 (1992).*
2. R.Kronstrand. Identification of N-methyl-1-(3,4-methylenedioxyphenyl)-2-butanamine (MBDB) in urine from drug users. *J. Anal Toxicol. **20**: 512-516 (1996).*
3. M.S. Kaufman. *Tutorial.* Mass spectrometric fragmentation lecture for Armstrong Drug Testing Division.
4. D.J.Wells. Drugs-of-abuse testing. *Advance/Laboratory. **6**: 39-43 (1997).*
5. R.C. Clark, J.DeRuiter, F.T. Noggle jr., A. Valaer. Identification of 1-(3,4-methylenedioxyphenyl)-2-butanamines related to MDMA. *MICROGRAM. xxv111:154-168 (1995).*
6. Armstrong Laboratory Drug Testing Division. Procedure for extraction of amphetamines by solid phase and liquid - liquid Extraction.January 31st(1994).
7. M.E. Bronson, W. Jiang, J. DeRuiter, C.R. Clark. Structure - activity relationships of BDB and its monoethyl and dimethyl derivatives. *Pharmacol- Biochem - Beh. **51**:477-479.*
8. S.R. Needleman, R.W. Romberg. Limits of linearity and detection for some drugs of abuse. *J. Anal. Toxicol. **14**: 34-38 (1990).*

A WAY TO CONDENSE THE TIME CONSUMING PROCEDURE OF COGNITIVE TASK ANALYSIS

JoAnne E. Bates
Graduate Student
Department of Psychology

University of North Dakota
Grand Forks, ND 58202

Final Report for:
Summer Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

A WAY TO CONDENSE THE TIME CONSUMING PROCEDURE OF COGNITIVE TASK ANALYSIS

JoAnne E. Bates
Graduate Student
Department of Psychology
University of North Dakota

Abstract

Cognitive task analysis is a complex, time depleting process. Cognitive task analysis (CTA) is a group of processes that are aimed to designate the knowledge, skills and organization of fulfilling a certain task. This report will summarize the work that I contributed to the ongoing, substantial cognitive task analysis project named DNA (Decompose, Network, Assess). A brief description of the DNA project will be given, as well as short explanations of literature research conducted and specific tasks in which I participated with a goal of furthering the DNA project.

A WAY TO CONDENSE THE TIME CONSUMING
PROCEDURE OF COGNITIVE TASK ANALYSIS

JoAnne E. Bates

Introduction

Shute and Torreano (1995) proposed a standardized cognitive task analysis procedure titled DNA (Decompose, Network, Assess) that could be applied across different domains. Shute, Willis and Surgue (1997) have cultivated the beginnings of the DNA process and developed a prototype to test. I had the opportunity during my summer internship at Armstrong Laboratory (AL/HRCT) to learn about the DNA project and to contribute to the continuing development of it. This paper will (1) briefly summarize the workings of the DNA scheme, (2) explain some examples of specific issues that I worked on to contribute to the DNA procedure, and (3) include short review about implicit learning that was done to see how the findings relate to the DNA procedure.

DNA Procedure

DNA (decompose, network, assess) is intended to obtain information from subject-matter experts (SME) by decomposing a domain, then network the information into extensive graphical structures (nodes and links), and then utilize other experts in that domain to assess the validity, totality, and reliability of the information structures. The information obtained by DNA will then be used by an instructional designer (ID) to set up an intelligent tutoring program to teach whatever domain was decomposed, networked and assessed. Usually the process of obtaining the domain information (CTA) is done through interviews and research. The DNA system is designed to automatize CTA and cut the time it takes to perform CTA.

DNA is set up into five modules. Module one is the customize module and this module is completed by the instructional designer (ID)

and will be given to the SME to install on their his/her computer to be able to run the decompose module. The ID will set up guidelines for the SME as to what domain to cover and what restraints should be put on the domain information due to the learning. Module two is the orientation modules which contains instructions on how to use DNA and explains What, How, and Why questions. What, How and Why questions are used by DNA to derive symbolic knowledge (e.g., What is the definition of…?), procedural skills (e.g., How do you…?), and conceptual knowledge (e.g., Why is that important?). The third module is the decompose module, which the SME uses on his/her PC, and includes a organization of set questions that the program poses to the SME. This is an intricate process that is designed, by using what, when and how questions, to extract information from the SME about a domain. The forth module, the network module, uses the information gained by the decompose module and sets it up into nodes so the SME can then graphically arrange the information. Then once the SME networks the information about a domain out graphically the ID can use this to assess the validity of the information. This is the assess module of the DNA procedure, the fifth module. The ID takes the networked information and compares it to other SMEs' graphical networks and also gets other SMEs to validate or change/add to networks from other SMEs (Shute, Willis, & Sugrue, 1997). The assess module is a verification process so then the ID can make sure the information is appropriate to then design a tutorial program using SMART: Student modeling Approach for Responsive Tutoring (see Shute, 1995, for details).

I was fortunate to get to work on a few aspects of the decompose module of DNA. I aided in finding a list of frequently used words to add to the dictionary of DNA so that DNA would not ask SMEs to define

common words. I worked out domain scenarios on the decompose module and worked on examples and pointers for the decompose interface.

Following are some examples that we worked on to help explain how the decompose module runs. These examples are on the interface of the decompose program and the SME, while using the program, can use these examples to help guide there answering about a domain. These examples are comprised of what (symbolic knowledge), how (procedural knowledge), and why (conceptual knowledge).

String of Questions and Answers for Example Interface on Decompose module of DNA:

Example 1 - Baking a pie - DNA: What can you tell me about baking a pie? SME: I know how to make pie crust. DNA: What do you generally do first when you make pie crust? SME: Mix the ingredients (1/3 cup butter, 2 cups flour, 1 teaspoon salt, 1/3 cup ice water) for the pie crust dough together in a bowl. DNA: What is the definition of pie crust dough? SME: A batter that when baked turns to a flaky pastry-like shell to hold pie ingredients. DNA: Why do you make pie crust? Because you need to have the crust (shell) to hold the filling ingredients. DNA: What is a typical instance when you make pie crust? SME: When I want to make a pie for dessert. DNA: What makes you good at baking a pie? SME: Following the order of the steps in the recipe.

DNA: What else can you tell me about baking a pie? SME: I understand that the same pie crust recipe can be used for many pie fillings. DNA: What is the definition for pie fillings? SME: Ingredients that you put on top of the pie such as cherry filling, apple filling, pecan filling. DNA: Why is it important that you understand that? SME: Because then you can perfect the same pie crust recipe and make cherry pie, apple pie, pecan pie, etc. and quiche. DNA: How can you show me that you

understand? SME: I can use the same recipe to make many different kinds of pies and quiches.

Example 2 - laundry-DNA: What can you tell me about doing laundry? SME: I know how to sort clothes. DNA: What do you generally do first when you sort clothes? SME: Separate darker colored clothes from white or light colored clothes. DNA: Why do you separate clothes? SME: To prevent the colored clothes from bleeding onto the white or light colored ones. DNA: Define bleeding. SME: transfer of dye from one cloth to another. DNA: What is a typical situation in which you sort clothes? SME: When I have colored and white clothes that need washing. DNA: What makes you good at doing laundry? SME: Following the washing instructions on the washing machine, on the clothes and on the detergent box. DNA: What else do you know about doing laundry? SME: I understand that different water temperatures (hot, warm, cold) effect different types of cloth. DNA: Why is it important that you understand that? SME: Because the different temperatures serve different purposes in washing clothes. For example, hot water washes very soiled clothes better than warm or cold, yet hot water makes colors bleed more and may shrink certain materials such as cotton. DNA: How can you show me that you understand? SME: I can wash several loads of clothes and use the correct temperatures for washing them and they will come out clean and just like new.

Example 3 - driving a car with standard transmission- DNA: What can you tell me about driving a car with standard transmission? SME: I know how to shift gears. DNA: What's the first step? SME: step 1 for shifting - depress clutch. DNA: define clutch. SME: the foot peddle to the left of the brake peddle. DNA: Why do you depress the clutch? SME: To prevent stripping the gears. DNA: Define stripping. SME: ruining the standard transmission in a car by not depressing the clutch before

moving the gear shift stick. DNA: What's a typical situation of depressing clutch? SME: to shift gears or brake to a stop. DNA: What else can you tell me about driving a manual car? SME:I understand that the clutch and gear shift stick are the manual equivalent to an automatic transmission in a car. DNA: Why is it important that you understand that? SME: In order to drive a standard transmission. DNA: How can you show me? SME: I can shift from first to second gear.

Example 3 - Diagnosing a common cold - DNA: What can you tell me about diagnosing a common cold? SME: I know how to take someone's temperature. DNA: What's the first step? SME: Step one - shake thermometer. DNA: define thermometer. SME: device used to measure body temperature. DNA: Why do you shake thermometer? SME: To move the temperature indicator to the bottom of the scale (below 95 degrees). DNA: What's a typical situation when you would do this? SME: When you suspect someone might have a fever. DNA: What else can you tell me about dianosing a common cold? SME: I understand that cold symptoms can be confused with symptoms of other health problems. DNA: Why is this important? SME: Because cold symptoms (coughing, stuffy nose, runny nose, sore throat, etc.) could be symptoms of a more serious disease. DNA: How can you show me that you understand this? SME: I can tell you that people that are diagnosed with lung cancer may have beginning symptoms that are similar to cold symptoms.

One idea that should be explored when trying to thoroughly extract information or knowledge from an SME about a domain concerns explicit and implicit learning and knowledge. I reviewed current literature on implicit knowledge and learning to find any information that would aid in the development of the DNA modules

Implicit Learning

The term *implicit learning* was presented by Reber (1967) to signify a type of learning that is presumed to occur without intention to learn and without coexisting awareness of the fundamental structure of the to-be-learned material (Eimer & Goschke, 1996). Reber, Kassim, Lewis, and Cantor (1980) showed, while studying modes of learning a complex rule structure, that tasks are better undertaken with implicit experience if the acute aspects are hidden in a mass of less relevant data, however the situations that respond well to the explicit mode of learning are when the key aspects of the task are fairly important and distinct.

Individuals are able to learn complex rules of system behavior even when there is no conscious attempt to do so. This classification of learning, referred to as *implicit learning,* is characterized as the acquisition of information regarding stimulus covariation acquired through successive presentation of problem examples without intent or cognizance. The information about a task is depicted by an conceptual rule or model that relates patterns of stimuli to particular responses. Implicit learning is ascertained by enhanced fulfillment of tasks that seem to demand thinking although the individual is not able to verbalize the rules utilized.

Implicit learning is nearly resource independent, which is different from explicit learning that relies on cognitive resources and is susceptible to distraction. Other terms have been used to designate this implicit/explicit learning phenomena, they are: incidental vs. intentional learning; selective vs. unselective learning; intuitive vs. deductive problem solving; and factual vs. tacit knowledge (DeShon & Alexander, 1996).

Experimental prototypes used to test implicit learning are probability learning, covariations learning, and learning complex rule

systems. It has been implied that perceptual knowledge (vs. symbolic processing) is fundamentally responsible for implicit learning and that implicit learning is linked to exterior characteristics.

Implicit memory studies focus on individuals' unconscious retrieval of a pre-made depiction, while implicit learning studies emphasize the implicit obtainment of fresh information. Nonetheless, both types of studies precede to a similar conclusion: the implicit system is more robust in that it displays more opposition to both cognitive and psychiatric disorders than the explicit system does. Implicit knowledge is harder to modify and more enduring, while explicit knowledge is more malleable and short-lived.

Implicit learning generates "abstract but possible instantiated implicit knowledge, although the exact representation of this knowledge is unclear… Implicit learning may involve changes in high-level perceptual representation systems in a manner similar to priming" (Seger, 1994).

Whittlesea and Wright (1997) determined that individuals generally seem to learn passively in implicit-learning experiments, reacting to only the conventional organization of the stimuli because the task is held constant within the experiment and because the stimuli include no other information to respond to. They state that these are the reasons individuals respond as they do, not because the learning system is innately passive or stimulus driven. They say that there is a wide agreement that humans have two alternative modes of learning, one implicit and one explicit, that follow dissimilar rules under differing circumstances. These two modes are said to contrast in many ways, but probably most significantly on the dimension of control. Implicit learning is controlled fundamentally or solely by changes in the stimulus circumstances, while explicit learning is controlled from

within the individual by consciousness and planning. However, these authors think that this idea manifests a false dichotomy. They maintain that there is just learning, operated by only one set of principles, and that in each instance of learning, some aspects will be explicit and some implicit.

Though Whittlesea's and Wright's (1997) ideas on the nature of the combination of implicit and explicit learning is reasonable and useful, there is research that supports the idea of two alternative modes of processing, each with its own advantages.

How can implicit and explicit processing best be represented? Braodbent, FitzGerald, and Broadbent (1986) indicated that there are two modes of processing, one for implicit and one for explicit. Implicit processing takes place by means of a "look-up table" in which situations and fitting actions are stored and checked with as needed. Access to the look-up table is quicker and more appropriate for real-time situations. Explicit processing is best conceived of as a model manipulation, or a "look-ahead strategy." Look-ahead strategy entails a serial process. The steps of this serial process are verbally reportable. However, in accordance with Whittlesea and Wright (1997), Broadbent et al. (1986) asserted that both processing modes may be operating in the performance of a single task.

Reber (1985) also agreed that knowledge gained from a simple observation of stimuli is neither totally implicit or explicit, but lies on a point along some "implicit-explicit continuum."

A substantial question throughout the literature is: Can you convert implicit knowledge to explicit knowledge? That's a good question for which there are no direct answers, yet. It has been shown that, within implicit learning research, the conceptual fluency response modality evaluates implicit knowledge and makes explicit judgments based

on it. However, the underlying knowledge remains implicit. There is evidence that implicit learning results in representations that are not flexible (e.g., people do not have access to components of the knowledge that they can use for other purposes). The information learned my be linked to the responses made in the task and cannot be transferred to other responses. Implicit learning may also be inflexible because of its nonhippocampal base. It is argued that the hippocampus is needed for flexible knowledge representation.

If implicit knowledge could be converted to explicit knowledge, then the knowledge could be used in a wider range of contexts. Possible development implicit and explicit knowledge together could facilitate then redescribing implicit representations into explicit representations. There are a few theoretical models that have been hypothesized in developmental psychology of how infants and children innately redescribe implicit knowledge to become more flexible, but these are not very useful for understanding how adult experts can redescribe their implicit knowledge into explicit knowledge.

The intent of this literature review was to see if the findings contribute to the design of the DNA prototype program. From an extensive review of the literature on implicit knowledge, I found no procedure to accomplish a "vulcan mind meld" for extracting implicit knowledge from domain experts. Converting implicit knowledge to explicit knowledge is something that psychology has not even come up with a good model for yet.

However, I feel that the research I reviewed does substantiate the idea that DNA should explain implicit knowledge and procedural automaticity to each expert that is performing the decompose module of DNA, as well as the experts that are validating the information on a domain. A strong emphasis introduced to the SME during the decompose

module and possible reminders throughout the program may help to ensure that DNA extracts all possible "forgotten" or "hard-to-reach" knowledge from each SME.

The nature of expertise is a broad topic, and in terms of DNA, the main idea I think should be focused on for DNA is that there are different levels and types of experts in each domain. This was considered when the Assessment part of DNA was created. Validating domain information from one expert by others is taking into account the idea that "many heads are better than one," or that experts vary in their expertise of a domain.

Conclusion

The prototype for DNA is progressing very quickly and it is an exciting project to research. The notion of automatizing and standardizing the CTA process is fairly new but the DNA procedure is adding fresh ideas to CTA. It will be very interesting to see what data is obtained when testing of the prototype begins. Positive results will come from experimentation with the DNA modules that will tell us more about CTA and knowledge in general.

References

Braodbent, D.E., FitzGerald, P., & Broadbent, M.H.P. (1986). Implicit and explicit knowledge in the control of complex systems. British Journal of Psychology, 77, 33-50.

DeShon, R.P., & Alexander, R.A. (1996). Goal setting effects on implicit and explicit learning of complex tasks. Organizational Behavior and Human Decision Processes, 65(1), 18-3.

Eimer, M., & Goschke, T. (1996). Explicit and implicit learning of event sequences: Evidence from event-related brain potentials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(4), 970-987.

Reber, A.S., Allen, R., & Regan, S. (1985). Syntactic learning and judgements: still unconscious and still abstract. Journal of Experimental Psychology: General, 114, 17-24.

Reber, A.S., Kassim, S.M., Lewis, S., & Cantor, S. (1980). On the relationship between implicit and explicit modes of learning a complex rule structure. Journal of Experimental Psychology: Human Learning and Performance, 6(5), 492-502.

Seger, C.A. (1994). Implicit learning. Psychological Bulletin, 115(2), 163-196.

Shute, V.J. (1995). SMART: Student modeling approach for responsive tutoring. User Modeling and User-Adaapted Interaction, 5, 1-44.

Shute, V.J., & Torreano, L.A. (1995). Cognitive task analysis made simple: Decompose, network, assess (DNA). Unpublished manuscript.

Shute, V.J., Willis, R.E., & Sugrue, B. (1997). Automating cognitive task analysis. Paper presented in: *Cognitive Technologies for Knowledge Assessment* symposium, AERA, Chicago, IL, March, 1997.

Turner, C.W., & Fischler, I.S. (1993). Speeded tests of implicit knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(5), 1165-1177.

Whittlesea, B.W.A., & Wright, R.L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. Journal of Experimental Psychology: Learning, Memory and Cognition, 23(1), 181-200.

EFFECTS OF BRAIN TEMPERATURE ON FATIGUE IN RATS DUE TO MAXIMAL EXERCISE AND MILLIMETER MICROWAVE RADIATION

Brandon B. Boke
Graduate Student

Trinity University
715 Stadium Dr.
San Antonio, TX 78212

Final Report for:
Summer Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Brooks Air Force Base, TX

and

Armstrong Laboratory

August 1997

EFFECTS OF BRAIN TEMPERATURE ON FATIGUE IN RATS DUE TO MAXIMAL EXERCISE AND 2.07 GHz MICROWAVE RADIATION

Brandon B. Boke
Graduate Student
Trinity University

Abstract

In previous experiments, it has been observed that fatigue in rats may be linked to hypothalamic temperatures (T_{hyp}) and that exercising rats become exhausted once the T_{hyp} has reached approximately 40.7-41.5°C. We attempted to prove the validity of this observation by subjecting a group of rats to two separate treatments, a microwave (MW) and a sham exposure, each followed by a run to exhaustion. Hypothalamic and rectal temperatures (T_{rec}) were measured continuously during their run. Each rat underwent both the microwave and the sham treatments. For the microwave treatment, the rat was placed in a Styrofoam restrainer and its T_{hyp} was heated with MW radiation up to 41.5°C. During the sham treatment, the rat was also restrained, but not exposed to any microwave radiation. There were significant differences in both pre-run T_{hyp} and T_{rec} between sham and microwave rats. However, there was no significant difference between the two treatments for either the peak T_{rec} or T_{hyp}. Although there existed no significant difference between the mean time to exhaustion of the sham and microwave groups, there was a significant difference between the heating rate and weight loss of the two groups. Because there was no significant difference between the mean peak T_{hyp}, we conclude that there exists a critical brain temperature which limits the ability to perform exercise.

EFFECTS OF BRAIN TEMPERATURE ON FATIGUE IN RATS DUE TO MAXIMAL EXERCISE AND MILLIMETER MICROWAVE RADIATION

Brandon B. Boke

Introduction

Although many different hypotheses have been proposed in an attempt to explain fatigue in mammals under hyperthermic conditions, the physiological mechanism underlying this effect remains unknown. While fatigue in thermoneutral conditions occurs as a result of hypoglycemia and muscle glycogen depletion, in hyperthermic conditions it occurs in the absence of those biological effects (2). It has been observed that exhaustion in exercising rats occurs when the hypothalamic temperature (T_{hyp}) reaches 40.7-41.5°C. (4). This suggests a possible link between T_{hyp} and fatigue in rats. In order to test this link, we performed a series of experiments using both exercise and millimeter microwave radiation (MW) to raise brain temperatures in rats above thermoneutrality. Because the performance ability in rats decreases after T_{hyp} has exceeded 40.7°C, we hypothesize that rats possess a centrally-mediated thermoregulatory response which reduces the large contribution of muscular work to the thermal load of the animal. This mechanism manifests itself in the form of exhaustion, but ultimately serves as a protective function to prevent the brain from reaching damaging levels. The purpose of the experiment is to test this hypothesis. Two different treatments were used in the investigation. One consisted of the rat being restrained and exposed to MW, causing the T_{hyp} to rise to an arbitrary temperature. This was immediately followed by a run to exhaustion on a treadmill with a thermally constant environment. The other treatment served as a control in which the rat was only restrained and then run to exhaustion. If our hypothesis proved to be correct, we would expect the rats to reach exhaustion faster following MW treatment because the T_{hyp} begins at a significantly higher temperature.

Methods

All rats were anesthetized with a combination of Ketamine (70 mg/kg; i.p.) and Xylazine (10 mg/kg; i.p.).

They were then placed on a thermostatically controlled water heating pad set to maintain a rectal temperature (T_{rec}) of 37°C. Stereotaxic surgery was then performed in order to implant Vialon guides (Becton Dickinson) into the hypothalamus of the rat brain. The coordinates for the tip of the hypothalamic guide were 1.8 mm posterior to bregma, 1.5 mm lateral from midline, and 8.3 mm below dura. The top of the hypothalamus guide was equipped with a threaded head cap connected to a Tygon TM tether tube used to contain the thermal probe. The guide was sealed at the bottom and held in place by cranioplastic cement (Plastic One, Roanoke, VA) anchored to nylon screws. These methods have been previously described in detail (3,5)

After two weeks of recovery from the surgery, the rats were familiarized with the Styrofoam restrainer and the motorized treadmill (Model: Dual Economy, Columbus Instruments, Columbus, OH). Each rat received four training sessions spaced no less than two days apart. Every session consisted of a period of familiarization in the restrainer and a ten minute run on the treadmill at a speed of 13.0 m/min. This protocol has been shown not to induce a training effect (1). The periods of familiarization began at 10 minutes and increased in duration up to 25 min.

Every rat in the experiment (n=10) was assigned a number, and each rat was eventually subjected to both treatments. In order to maintain a random design for the experiment two labels were made for all ten rats, one for sham (SH) and one for MW. These labels were then placed into a box and drawn one by one to determine the order in which the rats were used. After a rat had been selected for one treatment, the other label corresponding to that rat was removed until all the rats had been subjected to the first cycle of experiments. The rats were then subjected to the opposite treatment in the same order they were drawn to complete the second cycle of experiments.

All experiments were carried out between 1300-1500 hr. Rats to be used were restricted from feeding the morning before an experiment. Experiments were performed in an anechoic microwave chamber with an environmental temperature (T_{env}) of 21-23°C. Rats were brought into the chamber and weighed before each experiment. Post experiment weights were also measured to account for weight loss. Feces were

also collected from the end of the treadmill and subtracted from the total weight loss to yield fluid weight loss.

For all experiments, the rat was first fitted with a temperature probe inserted 5 cm beyond the rectum to measure rectal temperature (T_{rec}). The nonmetallic temperature probe (+/- 0.1 C, Model 101, Vitek, Boulder, CO) was secured by four pieces of silk surgical tape wrapped around the probe and the tail. With one individual operating the headgate, another experimenter inserted the rat into the rear of the Styrofoam restrainer. With the headgate closed over the rats' neck and a backstop in place behind the rat, the subject was then restricted from movement and ready for MW exposure. In the SH experiments, the rat was simply held in the restrainer for 5 minutes with no MW exposure.

All MW exposures (2.06 Ghz; CW) were conducted at an incident power density of 200 mW/cm^2 at the head with rats in the K, E, H polarization. During both MW and SH experiments, a Vitek probe was inserted into the hypothalamus guide and held in place by the Tygon TM tether and threaded screw. T_{hyp} was monitored at all times. Once the exposure began, T_{hyp} was allowed to rise to 41.5°C and then the MW transmitter was turned off.

Immediately following the MW or SH treatment, the rat was removed from the restrainer keeping both the rectal and hypothalamus probes attached to the rat. The rat was then placed on the treadmill already running at 17.0 m/min. The enclosed treadmill was on an 8% grade and was preheated to the target T_{env} of 35°C by a thermostatically controlled heating source. The T_{env} was maintained throughout the experiment within 1°C of the target T_{env}. In order to encourage the rats to run, the treadmill was also equipped with an electrical shocker on the end of the treadmill belt. The current running through the shocker was held constant throughout all the experiments.

Temperatures were manually recorded every five minutes beginning when the rat was placed upon the running treadmill. Vitek instrument displays and a profile of the running rat were also recorded on videotape for data backup and more detailed analysis at a later time.

Once the rat reached exhaustion, which was defined as the point in time at which the rat, despite electric shock, was no longer able to keep pace with the treadmill, it was removed and allowed to cool off on the

chamber floor. The point of exhaustion was determined by an investigator blind to both the temperature measurements and the exercise time. The rats were given a minimum of two weeks to recover from their first run before undergoing the other treatment.

Statistical analysis was performed using Statistica (Versiion 5.0, StatSoft, Tulsa, Ok) software, t-tests on dependent variables for repeated measures.

Results

During the MW treatments, the mean T_{hyp} following exposure was 40.8 ± 0.06°C. The mean T_{hyp} following the restraining period during the SH treatments was 39.3 ± 0.06°C. This represents a significant difference in pre-run hypothalamic temperatures between the MW and SH treatments (Fig. 1).

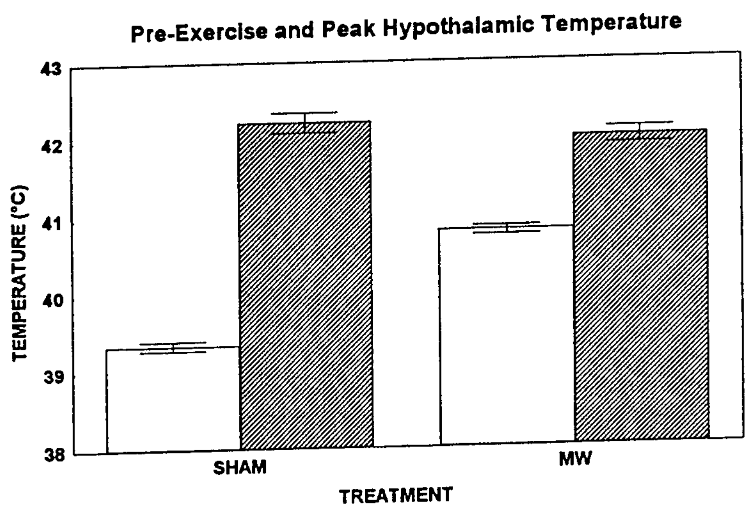


Figure 1.

There was also a significant difference in pre-run rectal temperatures between the two treatments with the MW T_{rec} at 40.4 ± 0.07°C and the SH T_{rec} at 39.4 ± 0.11°C (Fig. 2).

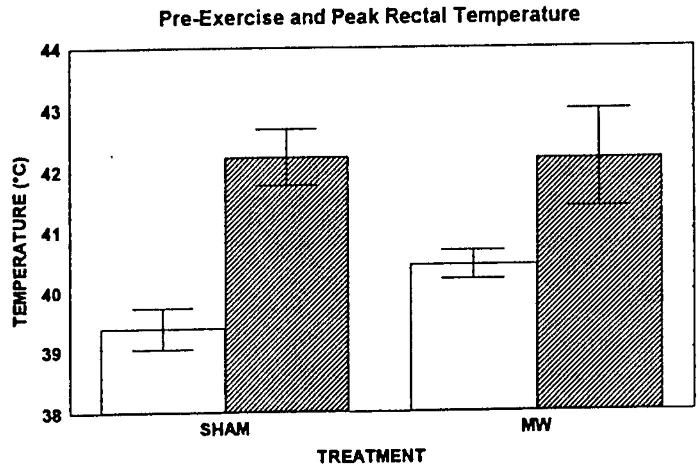


Figure 2.

Peak hypothalamic and rectal temperatures showed no significant difference between the MW and SH treatments. Peak MW T_{hyp} was 42.0 ± 0.11°C, and peak SH T_{hyp} was 42.2 ± 0.13°C (Fig. 1). Peak MW T_{rec} was 42.2 ± .16°C, and peak SH was 42.2 ± 0.27°C (Fig. 2).

The mean time to exhaustion for the MW treatment was 33.5 ± 4.8 min(Fig. 3). For the SH treatment, mean exhaustion time was 44.4 ± 4.8°C(Fig. 3). T-tests on this data showed no significant difference between the two treatments (p=0.067).

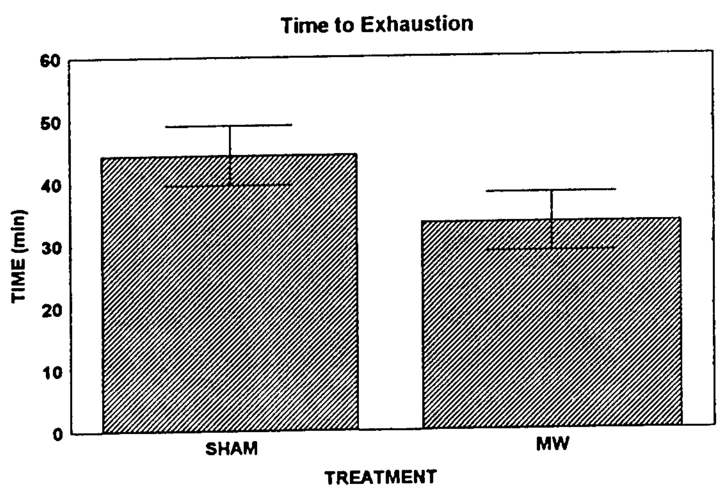


Figure 3.

A significant difference was found when comparing the heating rate of the rats between the two treatments. The mean heating rate was lower for the MW treatment at 0.040 ± 0.006 °C/min., where the rate was 0.073 ± 0.009°C/min. for the SH treatment (Fig. 4).

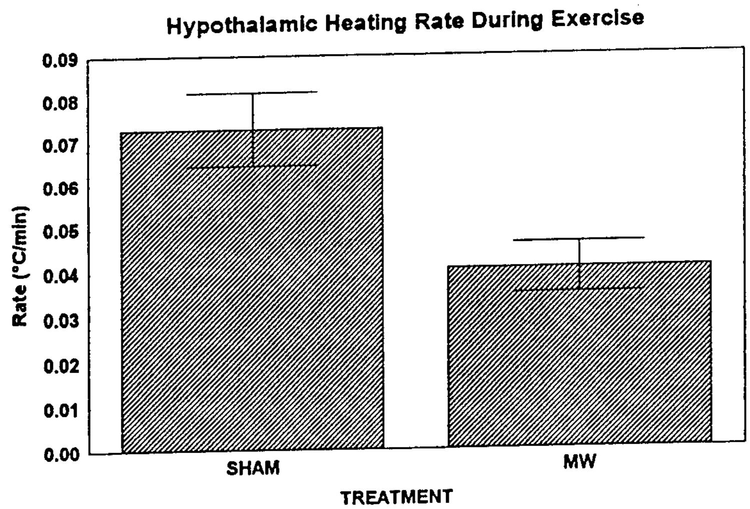


Figure 4.

T-tests looking at the mean weight lost during the runs also revealed a significant difference between the two treatments. During the MW experiments the rats lost an average of 5.0 ± 0.9g of weight. They lost an average of 6.3 ± 0.9g during the SH treatments (Fig. 5).

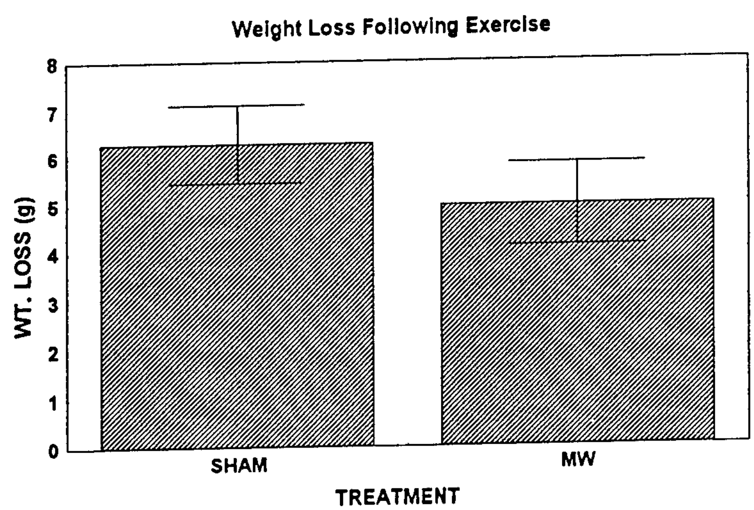


Figure 5.

Conclusion

As we expected, the peak hypothalamic temperatures for both groups showed no significant difference. This indicates that each individual rat has a critical brain temperature beyond which work capability is compromised. This is shown to be true regardless of how fast the animal heated up. The rate of heating could, however, explain why no significant difference was found between the exhaustion times of the two treatments. We expected the animal to reach exhaustion faster during the MW treatment because the rats began their run at a significantly higher temperature. This, most likely, would be the case had the animals' rate of heating been the same in both treatments. Since the rats heated up significantly slower during the MW treatments, it took longer for these rats to reach the critical T_{hyp} which causes exhaustion. There are a number physiological factors that are known to contribute to exhaustion during maximal exercise. Two of these, lactate acidosis and muscle glycogen depletion, are very important, but, unfortunately, measuring their effects is neither practical nor possible in a repeated measures study such as this. An indwelling catheter is required in order to measure lactate levels during exercise. Given the logistic complexity of the study as it is, this procedure would be render the study inoperable. In order to measure glycogen depletion, it is necessary to remove the muscle tissue from the rat so it can be ground and analyzed. Hypoglycemia is another factor to consider, but related pilot studies indicate it does not affect this experiment by significant measures.

One factor that we were able to measure and that could possibly have the most influence of them all is dehydration. By measuring the amount of fluid lost, either by evaporation, urine, or saliva, we found that the rats lost 6-14% of their plasma volume level. This is a significant amount of fluid loss and could greatly affect the time the rats reach exhaustion. Why there was a significantly greater amount of fluid loss during the SH treatments is a question yet unanswered. This and other questions will be addressed further in this ongoing study.

1. Cartee, G., and R. P. Farrar. Exercise training induces glycogen sparing during exercise by old rats. J. Appl. Physiol. 64: 259-265, 1988.
2. Holloszy, J.O., and W.M. Kohrt. Regulation of carbohydrate and fat metabolism during and after exercise. Annu. Rev. Nutr. 16:121-138, 1996.
3. Mason, P.A., T. J. Walters, J.M. Doyle, J.L. Kane, R. Escarciga, and K.E. Purdy. Heat-shock protein and GFAP expression after microwave exposure or warm water immersion. FASEB J. 9: A546, 1995.
4. Walters, T.J., K.L. Ryan, M.R. Tehrany, M.B. Jones, L.A. Paulus, and P.A. Mason. HSP70 expression in the CNS in response to exercise and heat stress in rats. J.Applied Physiol., Submitted.
5. Walters, T.J., K.L. Ryan, J.C. Belcher, J.M. Doyle, M.R. Tehrany, and P.A. Mason. Regional brain heating during microwave exposure (2.06-Ghz), warm-water immersion, environmental heating, and exercise. J. Thermal Biol., Submitted.

ASSESSMENT OF COAGULANT AGENTS ON THE REDUCTION OF AQUEOUS FILM FORMING FOAM (AFFF) IN WASTEWATER

Constance Buford
Graduate Student
Department of Plant & Soil Sciences

Alabama A&M University
P. O. Box 1208
Normal, Alabama 35762

Final Report for:
Summer Graduate Student Research Program
Armstrong Laboratory/Occupational Bioenviromental Engineering Water Quality Division (AL/OEBW), Brooks Air Force Base, Texas

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force, DC

and

Armstrong Laboratory

August 1997

ASSESSMENT OF COAGULANT AGENTS ON THE REDUCTION OF AQUEOUS FILM FORMING FOAM (AFFF) IN WASTEWATER

Constance Buford
Graduate Student
Department of Plant & Soil Sciences
Alabama A&M University

ABSTRACT

Aqueous film-forming foam (AFFF) concentrate is used by the Air Force (AF) to fight aviation fuel fires. Accidental spills of this material are collected in waste basins and are then transferred to wastewater treatment plants (WWTPs). The characteristics of the AFFF which make the compound a highly effective fire suppression system also make it toxic to aquatic animals and troublesome to wastewater treatment facilities due to its high BOD level and foam stabilizers. The use of coagulant agents to minimize the formation of foam was the central focus of this study. A detergent-Alconox, mineral oil, polymer - Chemlink, and defoamers (Calgon 926 and 935; Nalco 2508-Plus) were used to treat a 1L sample of 300 mg/kg concentration of 6% AFFF. The defoamer 2508-Plus by Nalco was found to be the best coagulant. The 300 mg/kg concentration of 6% AFFF, once treated with the Nalco product and after additional agitation, did not re-foam. The optimum dosage for 300 mg/kg concentration of 6% AFFF is 1.75 mls or 35 mg/kg of 2% Nalco 2508-Plus. The treatment cost would be $0.01 per 1.75 mls. On a larger scale of about 40,000 gal 300 mg/kg of 6% AFFF the treatment costs would be @ $1778.15. Continued study, however, to identify the most effective coagulant with the least expensive dose of chemicals will provide a more cost-effective method for the treatment of AFFF in wastewater.

ASSESSMENT OF COAGULANT AGENTS ON THE REDUCTION OF AQUEOUS FILM FORMING FOAM (AFFF) IN WASTEWATER

Constance Buford

INTRODUCTION

Aqueous film-forming foam (AFFF) concentrate, conforming to MIL-F-24385, is used by the Air Force (AF) to fight aviation fuel fires. AFFF contains water, ionic and non-ionic surfactants (fluoroaliphitic detergent), other additives (organic surfactants), and diethylene glycol monobutyl ether (or butyl carbitol), a foam stabilizer which is toxic to aquatic animals and is troublesome to water treatment facilities. The surfactant components of AFFF mixture, even in small concentrations, results in significant foaming which severely impacts the operations of waste water treatment plants (WWTPs). AFFF cannot be allowed to enter a sewer without a period of decomposition and cannot be allowed to enter surface or groundwater. Currently, there is no technology to effectively treat water contaminated by AFFF fire suppressant. Conversely, the procedure used consists of containing wastewater contaminated with AFFF in holding ponds (up to 20 days) followed by slowly feeding small amounts (< 5 gal.) of wastewater to the local WWTP over very long periods of time. (1, 5, 7, 8)

BACKGROUND

AFFF is used in a standard 1, 3, or 6% solution in the Deployable Fire Protection System (DFPS) (9). The triggering systems utilize infrared and ultraviolet (UV) light sensors (1). AFFF actuation can be caused by backfires or power supply problems (5). The problem with AFFF is twofold: When it is used to fight fires, the runoff must be collected and ponded if possible. When there is an accidental spill, the spill must be cleaned up and ponded. An accidental spill of the material is a reportable spill under the Clean Air Act (9).

The Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) requires reporting of any spill that exceeds the reportable quantity (RQ). This requirement includes reporting spills of "glycol ether" which are currently listed as a

hazardous air pollutant in Title III, Section 112, of the Clean Air Act. AFFF contains diethylene glycol monobutyl ether, which is considered a glycol ether under this act. Spills of glycol ether which exceed one pound must be reported to the EPA national response center. However the use of AFFF in fire fighting is not a "spill", spills in this case are accidental releases. Since AFFF is used in systems which use fire detection sensors and these sensors are fallible and accidental, reportable spills occur. (1,5, 7, 8, 9)

PURPOSE

This study addresses a means to alleviate problems associated with the release of AFFF from the flooding hangars as a fire suppressant. The purpose of this study was to suppress the amount of foam that is being produced.

OBJECTIVE

1) To determine the most effective coagulant and dosage in the treatment of AFFF.

LITERATURE REVIEW

The concern for water quality, especially with use of AFFF, has in recent years accelerated the demand for new treatment technologies for wastewater containing AFFF. Among the modifications to conventional treatments that can improve removal of AFFF, coagulation and flocculation are the most attractive.

Coagulation and flocculation involve the addition of chemical coagulants such as aluminum sulfate, ferric chloride, and/or polyelectrolytes (polymers) to raw water to hasten the settling of suspended matter. Since coagulation is not an exact science, selection and optimum dosages of coagulants are determined experimentally by the jar test, instead of quantitatively by formula. (10, 11)

Jar Testing

Jar testing has been commonly used for many years (Figure 1). It is the quickest and most economical way to obtain good reliable data on the many variables which affect

the treatment process. However, there is no standard procedure for conducting the jar test, nor is there standard equipment that must be used. The jar test is performed using a series of glass container that hold at least 1L and are of uniform size and shape. Normally, six jars are used with a stirring device that simultaneously mixes the contents of each jar with uniform power input. Each of the jars contains a 1L sample of raw water and progressively larger doses of the coagulant chemical. Visual observation of the floc, along with turbidity readings of the settled water in jars, provides the data necessary to make determinations. Plant operations is most efficient when the lowest turbidity is obtained in finished water with the lowest cost for coagulant chemicals. (3, 10, 11, 16)

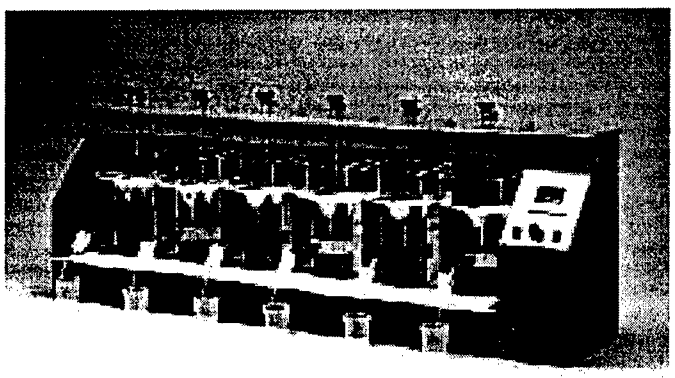


Figure 1. Phipps & Bird-700™ Series Six-Paddle Standard Jar Tester.

Wastewater Analyses

Certain wastewater characteristics affects the chemical selection of coagulants (11). The following should be considered:

- water temperature
- pH
- alkalinity
- turbidity
- conductivity

Water Temperature

Low-temperature water usually causes poorer coagulation and flocculation and can require that more of a chemical be used to maintain acceptable results. (11)

pH

The pH of water is a measure of its acidity or alkalinity. A scale of 0 to 14 is used for measurement with 0 being extremely acidic and 14 being extremely alkaline. The midpoint, 7, is neutral. Extreme values of pH, either high or low, can interfere with coagulation and flocculation. The optimal pH varies depending on the coagulant used. (10, 11)

Alkalinity

Alkalinity of a water is its acid-neutralizing capacity. It is the sum of all the titratable bases. The measured value may vary significantly with the end-pH used. Alum and ferric sulfate interact with the carbonates, bicarbonates and hydroxides that causes alkalinity in the water; thus reducing the alkalinity and forming complex aluminum or iron hydroxides that begin the coagulation process. Low alkalinity limit this reaction and results in poor coagulation; in these cases, it may be necessary to increase the alkalinity of the water. (2, 11)

Turbidity

Turbidity is caused by particles suspended in water. These particles scatter or reflect light rays directed at the water, making the water appear cloudy. Waters showing very little light scattering produce low turbidity measurements, whereas those with a great deal of light scattering indicate high turbidity. The lower the turbidity, the more difficult it is to form a proper floc. Fewer particles mean fewer random collisions and hence fewer chances for floc to accumulate. If this occurs, it may be necessary to add a weighting agent such as clay (bentonite) to low-turbidity water. (2, 11)

Electrical Conductivity

Electrical conductivity (EC) is a measure of the ionic strength of a solution and an indirect measure of the total suspended solids in a water (11). In general, every 10 units of EC represents 6 to7 mg/kg of dissolved solids (10).

Total Dissolved Solids

Total dissolved solids (TDS), also referred to as total filterable residue, are any particles of organic or inorganic matter that are dissolved in water — such as salts, chemicals, or gases. They are nonsettleable and can cause public health or aesthetic problems if not removed. Unless converted to a precipitate by chemical or physical means, they cannot be removed from the water. For these reasons, a limit of 500 mg dissolved solids/L is desirable for drinking waters. (2, 10)

APPROACH and METHODS

This study was conducted at the Armstrong Laboratory/Occupational Bioenviromental Engineering Water Quality Division (AL/OEBW), Brooks Air Force Base, Texas over a six week period. Jar testing was conducted on a 1L sample of 6% AFFF concentrate, in 1L glass containers, to identify best coagulant and dosage as described by Phipps and Bird, 1997. In addition to jar testing, the sample was analyzed for electrical conductivity, alkalinity, turbidity, and total dissolved solids (2).

The products used as coagulant agents in this study were: a detergent, mineral oil, polymers, and foam dispersing agents. The detergent was Alconox; general heavy mineral oil; the polymer was from Chemlink; the foam dispersing agents were from Calgon and Nalco. Cost of these products can be seen in Table 1.

Table 1. Cost of Disperseing Agents Per Pound.

Products	Weight	Cost
Chemlink	1 lb	$280.00
Nalco 2508 Plus	1 lb	$3.74
Alconox	1 lb	$3.25
Calgon 935	1 lb	$2.79
Calgon 926	1 lb	$2.16
Mineral OIL	1 lb	$1.92

DESCRIPTION OF FOAM AND DISPERSANTS

3M 6% AFFF

The 3M FC-206CF LIGHT WATER brand AFFF consists of 78-81% water; 9.5-10.5% of butyl carbitol [or 2-(2-butoxyethoxy)-ethanol]; 3-7% of urea; 1-5% of akyl sulfate salts; 1-2% of amphoteric fluoroalkylamide derivative; .1-1% of perfluoroalkyl sulfonate salts; .1-1% of triethonolamine; and 0-.1% of methyl-1H benzotriazole. It is a clear, amber liquid. It has a pH of 8.5 and a specific gravity of one. The biodegradation rate of this product is: 5-day BOD is 0.16g/g; 10-day BOD is 0.17g/g; 20-day BOD is 0.18g/g; and for COD, 0.34g/g (12).

DISPERSANTS

The chemicals in the dispersing agent are designed to weaken the bubble film and causes small bubbles to coalesce into larger ones that rupture easily at the liquid surface. Spreading agents help disperse the antifoam quickly and evenly through the foaming liquid. Experiments conducted by 3M Environmental Laboratory have shown that dispersing agents are effective in activated sludge or AFFF solutions for up to 3 hours. Once suppressed, the AFFF could re-foam if the liquid was agitated or aerated. Low shear pumps are recommended by the 3M manufacturer for pumping the AFFF without causing the product to re-foam.

ALCONOX

Alconox is a concentrated biodegradable detergent. It contains phosphates-7.3% by weight with a pH of 9.0-9.5.

MINERAL OIL

Mineral oil, the least expensive, is a tasteless, odorless, and colorless liquid.

NALCO 2508 PLUS

Nalco polymer is a clear, light yellow liquid with a distinctive odor. It is a blend of polyglycol, polyglycol ester, and fatty acid that work to control foam over a broad pH range. It is dispersible in water but will precipitate below 50^{o}F, and will solidify below 35^{o}F. At 77^{o}F it has a specific gravity of 0.84. (14)

CHEMLINK IPC 6850 ADDITIVE

Chemlink polymer, the most expensive product tested, is a white granular solid with a slight ammoniacal odor. The chemical's identity is a trade secret and is not available for public knowledge. It is partially soluble in water and has a specific gravity of one. The pH and percent organic compounds are unknown. (15)

CALGON DEFOAMER 926

Defoamer 926 is greenish brown, oily liquid with mild hydrocarbon odor. It is a hydrocarbon oil with silicone and amorphous gel. It is a severely refined petroleum distillate; silica, amorphous, precipitated and gel. It is not soluble in water and has a specific gravity of 0.90. (13)

CALGON DEFOAMER 935

Defoamer 935 is an opaque, white liquid. It is a blend of fatty alcohols, waxes, surfactants and mineral oils in water. It is a petroleum distillates; solvent-refined light paraffinic. It is dispersible in water. At 25^{o}C, it has a specific gravity of 0.92-0.98 and a pH of 7.5-10.0. (13)

RESULTS

Jar testing for coagulation of a 1L sample of 300 mg/kg concentration of 6% AFFF with six different types of foam dispersing agents was analyzed (Table 2). Since the amount of foam could not be quantified, the results are reported as low (minimal foam dispersal), and high (maximum foam dispersal). The 300 mg/kg concentration of 6% AFFF had a pH of 7.4; an electrical conductivity of 0.52 mS/cm; a turbidity of 0.92 NTU; total dissolved solids was 0.26; and an alkalinity of 20 mg/kg $CaCO_3$. The AFFF foams with the slightest agitation, but settled to its liquid form within thirty minutes.

Although the Mineral Oil is the least expensive product, as shown in Table 1, it was not the most effective on AFFF. However, by increasing the alkalinity of the 6% AFFF, the AFFF did not foam as much. The Alconox detergent and Chemlink polymer did not have an effect on the 300 mg/kg concentration of 6% AFFF. The defoamers, Calgon and Nalco, were more effective in reducing the foaming, yet Nalco was the most effective in dispersing the foam. Additional information from the manufacturers of the Nalco product shows that it should be applied at a concentration ranging from 5-100 mg/kg fed at 2-40% emulsion weight. Based on an initial release of 300 mg/kg of 6% AFFF, this volume could be treated with 1.75 mls of 2% Nalco defoaming agent. After testing the Nalco product, agitation of the AFFF did not induce foaming action for 72 hours.

Table 2. Effect of Coagulants on The Dispersion of 300 mg/kg of 6% AFFF.

Products	mls of coagulant used	Foam Dispersion
Chemlink	5-30	low*
Nalco 2508 Plus	0.25-1.75***	high**
Alconox	5-30	low
Calgon 935	0.35-0.60	high
Calgon 926	0.35-0.60	high
Mineral Oil	5-30	low

* minimal foam dispersal after 30 minutes

** maximum foam dispersal after 30 minutes

*** 1.75 mls had the most effect on the 300 mg/kg 6% AFFF

CONCLUSION

Jar testing was conducted to estimate the most effective coagulant and optimum dosage in the treatment of 6% AFFF at a concentration of 300 mg/kg. A detergent, mineral oil, polymers, and defoamers were used to treat the AFFF. The defoamer 2508-Plus by Nalco was found to be the best coagulant. The 300 mg/kg concentration of 6% AFFF, once treated with the Nalco product and after additional agitation, did not re-foam for 72 hours. The optimum dosage for 300 mg/kg concentration of 6% AFFF is 1.75 mls or 35 mg/kg of 2% Nalco 2508-Plus. The treatment cost would be $0.01 per 1.75 mls. On a larger scale of about 40,000 gal 300 mg/kg of 6% AFFF the treatment costs would be @ $1778.15. Studies, however, should be continued to better determine optimum dosages for individual plants. Also, supplementary studies on the biodegradation of AFFF and the coagulant or dispersant used to treat the AFFF should be continued.

REFERENCES

1. EA Engineering, Science, and Technology. April 1995. Fire-Fighting Foam Disposal Compliance Plan Andrew Air Force Base-Final. Sparks, Maryland.
2. Greenberg, Arnold E., Lenore S. Clesceri and Andrew D. Eaton. 1992. Standard Methods: For The Examination of Water and Wastewater 18th ed., American Public Health Association, Washington, DC.
3. Howard, S. Peavy, Donald R. Rowe and George Tchobanoglous. 1985. Environmental Engineering. pp. 136-140. McGraw Hill, Inc.
4. http://www.ehs.eduhsd.k12.ca.us/grflat/wtrqual 1.html
5. http://www.glue.umd.edu/~qureshi/andrews.htm
6. http://www.state.ky.us/nrepc/water/wcpcl.html
7. http://xre22.brooks.af.mil/96stratPlan/1A1.htm
8. http://xre22.brooks.af.mil/96TNS/96Needs/1609.htm
9. http://xre22.brooks.af.mil/96TNS/96Needs/971.htm
10. Huben, Harry Von (Sr. ed.). 1995. Water Quality 2nd ed. pp. 121-160, American Water Works Association, Denver, CO.
11. Huben, Harry Von (Sr. ed.). 1995. Water Treatment 2nd ed. pp. 51-83, American Water Works Association, Denver, CO.
12. Material Safety Data Sheet, p. 867, 3M Center, St. Paul, MN.
13. Material Safety Data Sheet, Calgon Corporation, Pittsburgh, PA.
14. Material Safety Data Sheet, Nalco, Naperville, IL.
15. Material Safety Data Sheet, Baker Industrial Chemicals, Corpus Christi, TX.
16. Phipps &Bird. 1997. A Simplified Jar Test Procedure. Phipps & Bird, Richmond VA.

THE EFFECTS OF INDIVIDUAL DIFFERENCES AND TEAM PROCESSES ON TEAM MEMBER SCHEMA SIMILARITY AND TASK PERFORMANCE

Dawn D. Burnett
Graduate Student

Department of Psychology
Wright State University

Final Report for:
Graduate Student Research Program
Wright Patterson AFB
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

THE EFFECTS OF INDIVIDUAL DIFFERENCES AND TEAM PROCESSES ON

TEAM MEMBER SCHEMA SIMILARITY AND TASK PERFORMANCE

Dawn D. Burnett
Graduate Student
Department of Psychology, Wright State University

Joan R. Rentsch
Associate Professor of Psychology
Department of Psychology, Wright State University

Michael D. McNeese
Research Scientist
CFHI, Armstrong Laboratory

Kristen R. Bonnema
High School Student
Wayne High School

Abstract

The research reported here examined team member schema similarity, its predictors, and its relationship to team performance. This research has implications for applied Air Force problems such as UCAV (Unmanned Combat Air Vehicle) operation, transport command centers, information warfare, battlefield management, C^3I, and joint collaborative systems such as data walls, avatars, intelligent agents, and knowledge rooms.

One purpose of the present study was to test a portion of the Team Member Schema Similarity Model with modifications. It was hypothesized that team member individual differences would predict team process variables (team interaction process variables and group process variables) and team member schema similarity (TMSS). Team process variables and TMSS were hypothesized to predict team performance.

Data were collected in a laboratory setting. Forty-five two member teams attempted to solve a complex, ill-defined problem. Team members completed individual difference measures before working on the problem. After solving the problem, team members completed teamwork schema measures. TMSS was operationalized as schema agreement and schema accuracy. Team performance and team interaction processes were coded by raters. Preliminary results are reported that indicate some support for portions of the modified model. Future research and potential applications are discussed.

The authors acknowledge the contributions of Martin Anesgart, Joe Balas, Darcy M. Menard, and Laura J. Pape.

THE EFFECTS OF INDIVIDUAL DIFFERENCES AND TEAM PROCESSES ON TEAM MEMBER SCHEMA SIMILARITY AND TASK PERFORMANCE

Dawn D. Burnett
Joan R. Rentsch
Michael D. McNeese
Kristen R. Bonnema

Teams, crews, multi-operator units, and collaborative systems abound in the Air Force. Dr. Michael McNeese has developed a research program to study the socio-cognitive variables operating within these multi-person units. Because Air Force teams operate in environments that are characterized by ill-defined and emergent situations, McNeese's research program has incorporated theories of situational awareness, crew schemata, shared cognition, social construction of knowledge, and interpersonal interaction to understand the development of meanings necessary for effective functioning in these environments (e.g., Brown, Whitaker, Selvaraj, & McNeese, 1995; McNeese, Zaff, Citera, Brown, & Whitaker, 1995; McNeese, 1993; Nosek & McNeese, 1997; Wellens & McNeese, 1987; Young & McNeese, 1995). This type of research will produce knowledge that will be applicable to the development of joint collaborative systems technology within the Air Force. The research study reported in this document is part of this research program.

The research reported below focused on cognitive processes within teams. Specifically, the research examined team member schema similarity, its predictors, and its relationship to team performance. This research is relevant to co-located or distributed teams. It also has implications for applied Air Force problems such as UCAV (Unmanned Combat Air Vehicle) operation, transport command centers, information warfare, battlefield management, C^3I, and joint collaborative systems such as data walls, avatars (Wells & Hoffman, 1996), intelligent agents, and knowledge rooms.

For example, UAV operation requires that pilots coordinate with field controllers, tower personnel, and manned air vehicles in complex, dynamic, ill-defined situations. During the chaos of combat, the UCAV crew must be able to construct meaning quickly in these ambiguous situations. UAV crew members must be able to

understand each other and they must have shared situational awareness. This ability to understand one another and to develop shared situational awareness is developed, in part, through team member schema similarity. Thus, team member schema similarity (TMSS) will enhance the performance of these types of teams.

Below, the conceptual model used to guide the current research is described, relevant past research is reviewed, and a modified version of the model is presented. Then the present study is described.

<u>Team Member Schema Similarity Model</u>

The Team Member Schema Similarity Model (Rentsch & Hall, 1994) is shown in Figure 1. The critical variable in the model is team member schema similarity (TMSS). Team member schema similarity (TMSS) refers to the degree to which team members have similar team-related schemas. A schema is a complex knowledge structure that organizes new information and facilitates understanding (Poole, Gray, & Gioia, 1990).

Although team members may develop many team-related schemas (e.g., schemas of team members, schemas of the task), the focus of the present study is on teamwork schemas (Cannon-Bowers & Salas, 1990). Teamwork schemas contain knowledge and information regarding communicating about, evaluating, and compensating for teammates' performance (Cannon-Bowers & Salas, 1990). Teamwork schemas will guide team members' assumptions, expectations, and behavior regarding teamwork.

Rentsch and Hall (1994) extended the concept of coorientation (Poole & McPhee, 1983), which is typically studied with respect to attitudes, to the study of teamwork schemas. They suggested that TMSS refers to schema accuracy as well as schema agreement. Schema agreement among individuals exists when individuals have similar schemas. This type of schema similarity has received research attention in the past (e.g., Walsh, Henderson, & Deighton, 1988). Schema accuracy exists when individuals are able to describe another individuals' schema accurately. This type of schema similarity has received very little research attention. A critical or optimal level of TMSS, in terms of both accuracy and agreement, is hypothesized to enhance team effectiveness. Empirical research has revealed some support for this hypothesis (Rentsch, Pape, & Brickman, 1997; Rentsch, 1993).

Team members who have schema similarity have similar knowledge about teamwork and they organize this information in a similar way. Teamwork schema similarity is hypothesized to enhance team effectiveness

because similar teamwork schemas in terms of agreement and accuracy among team members will allow team members to interact efficiently and effectively. Team members with high TMSS will be able to anticipate, to facilitate, and to compensate for one another's behavior.

Communication among team members is also likely to be enhanced as team members' teamwork schemas become increasingly similar. Team members may be aware of the information required by each other and fully understand the information that is being communicated to each other. Moreover, team members are likely to anticipate and understand each other's actions due to teamwork schema similarity.

Researchers in the military realm have studied team cognitions as team mental models. They have examined the relationship between team cognitions and team performance (e.g., Cannon-Bowers, Salas, & Converse, 1993). Much of this research provides indirect evidence that team mental models predict team performance (Cannon-Bowers et al., 1993).

Most researchers have conceptualized team cognitions in terms of schema agreement. The accuracy component of TMSS has not been explored extensively to the authors' knowledge. Thus, one purpose of the present study was to test the roles of team member schema accuracy and agreement in the prediction of team effectiveness. Rentsch, Pape, and Brickman (1997) examined TMSS conceptualized as schema agreement and schema accuracy. Their results revealed that agreement and accuracy predicted team effectiveness significantly. They suggested that schema accuracy may be more important in predicting team effectiveness than schema agreement. They suggested that future research examine these relationships in detail.

Antecedents of TMSS. As shown in Figure 1, two antecedents of team-related schema similarity are team membership influences, such as person-environment fit, and schema related communications, such as communication that occurs during socialization processes. These antecedents are expected to regulate the degree of schema similarity among team members to an optimal level. It is assumed that an optimal level of schema similarity exists for any given team and that when schema content is of high quality, the optimal level of schema similarity will enhance team effectiveness maximally.

In the present study, we focused on a subset of team membership influences and a subset of schema related communications as predictors of TMSS. The team membership influences we examined were individual differences. The schema related communications in our study were team interaction process variables. Below, we elaborate these portions of the model.

Team membership influences as individual differences. Schneider (1987) hypothesized that individuals are attracted to organizations containing people who are similar to themselves (Attraction), organizations select individuals who are similar to others already in the organization (Selection), and individuals who gain entry into the organization, but who are significantly different from those already in the organization are predicted to leave the organization (Attrition). One outcome of the ASA process is excessive similarity among organizational members. In the extreme, Schneider hypothesized, organizations may fail due to excessive homogeneity among organizational members, because too much homogeneity among organizational members would limit organizational creativity.

Thus, this organizational theorist predicted that similarity among organizational members will be associated with organizational members thinking similarly. In other words, individuals who are similar in terms of individual difference characteristics are likely to have similar schemas. Extending this logic to teams implies that similarity among team members' individual difference characteristics may enhance TMSS. Thus, an optimal level of TMSS may be due, in part, to similarities among team members' individual difference characteristics (Rentsch & Hall, 1994).

Although similarity among team member individual differences characteristics may be related to TMSS, it may also be the case that the level of the individual difference characteristics existing within the team may also be related to TMSS. Pape (1997) hypothesized that high levels of teamwork schema accuracy and agreement would exist when perceivers were high on trust and perspective taking, and when targets were high on private self-consciousness[1]. She hypothesized that targets high on self-monitoring would have low levels of accuracy and agreement. Team members who are high and similar in terms of several individual difference variables, such as perspective taking, trust, and private self-consciousness, and who are low and similar on self-monitoring are likely to experience high TMSS. These variables were investigated in the present study.

Schema related communications as team interaction processes. Schema related communication processes include communication that occurs during training, team member socialization, and team member interactions. In the present study we examined schema communication processes that occur while team members interact to complete a task. Within the team context, there is evidence that team member interaction is a primary cause of schema similarity among team members (Bettenhausen & Murnighan, 1991; Gersick, 1989; Walsh, Henderson, & Deighton, 1988). Many researchers agree that groups require social construction of knowledge and metacognitive processes to solve problems successfully (e.g., Young & McNeese, 1995). However, the research has not addressed the nature of the interactions leading to socially constructed meaning. No definitive theory for studying these interaction processes has been developed. Therefore, another purpose of the present study was to determine team interaction process variables that predict TMSS and team performance. We drew upon the group communication and crew coordination literatures. Our summary of this literature is presented next.

Team Interaction Processes

As Tower and Elliott (1996) note, although team member communication and coordination are hypothesized to be related to team performance, very little empirical research has been conducted to test the nature of these linkages. However, according to Tannenbaum, Beard, and Salas (1992), existing research evidence tends to support communication processes, coordination processes, and decision making processes as predictors of team performance.

For example, Hakel, Weil, and Hakel (1988) reported that intrateam communication patterns differentiated high from low performing teams. Among other findings, Kimble and McNeese (1987), found that teams that talked in longer utterances performed better than teams that spoke in shorter utterances.

Team interaction processes and TMSS are likely to have a reciprocal relationship. Team member communication is important because it is critical to developing similar understandings, and similar understandings, in turn, imbue communication with meaning (Fischer, 1996). Again, the research evidence is indirect. For example, Dyer (1984) reported research by Obermayer and colleagues revealing that experienced aircrews communicated more frequently during nonroutine missions than during routine missions compared to less

experienced crews. This suggests that experienced aircrews might have had TMSS for routine situations and, therefore, did not require high levels of communication for these situations.

Although there are many approaches for studying team interaction processes, there is none that specifically addresses the sense-making, or socially constructed aspect of interaction. Interaction seems to enhance performance and it seems to be related to TMSS. Yet, there exists a need for additional study of interactions and communication within teams (Duffy, 1993).

Team Interaction Process Variables

We developed a coding scheme to record the nature of the interactions occurring among team members. In particular, we focused on the degree to which team members tended to share interpretive information with one another. We assessed: depth of meaning of informing oriented communication, depth of meaning of learning oriented communication, quality of exchange, consensus evaluation communication, openness and acceptance of communication, and egocentricity versus mutuality of interaction. We also included group process variables in the present study. The group process variables were team cohesion, task motivation, and team metacognition.

The Present Study

The primary purpose of the present study was to test a portion of the TMSS model with modifications. The variables explored in the present study are presented in Figure 2. It was hypothesized that team member individual differences will predict team process variables (team interaction process variables and group process variables) and TMSS. Team process variables and TMSS are hypothesized to predict team performance.

Method

Participants

The participants were 90 undergraduate students who comprised 45 two member same-sex teams. There were 20 male teams and 25 female teams.

Task

The task used for the present study was *The Adventures of Jasper Woodbury: Rescue at Boone's Meadow.* The Jasper task is a complex, ill-defined task that requires problem solvers to identify the problem and the

subproblems needed to solve the problem, to distinguish between relevant and ill-relevant information, to coordinate relevant information, and to evaluate alternative possible solutions. Middle-school level mathematics are required to solve the problem (McNeese, 1993; Vye, Goldman, Voss, Hmelo, & Williams, undated).

The Jasper task is presented on laser video disc. A story problem is presented in which an injured eagle must be rescued from a remote location. Alternative routes, modes of transportation, and individuals are available for rescuing the eagle. Two problems are to be solved: (1) What is the quickest way to move the eagle? (2) How long will that take?

When solving the problem, team members could refer back to the video disc by using a Macintosh computer system. They were to record their answer and any information used to arrive at their answer in writing. The teams were videotaped as they attempted to solve the problem. They were stopped after 60 minutes, if they had not reached a solution. Only three teams were stopped. All other teams reached a solution in 12 to 56 minutes.

Measures

Individual differences. Traditional individual difference measures were used. The Interpersonal Trust Scale (Rotter, 1967) was used to assess trust. It had an internal consistency reliability estimate of .63 in the present study. In past research internal consistency reliability has been estimated at .76 (Rotter, 1967). Perspective taking was measured using the Interpersonal Reactivity Index (Davis, 1980), which is a 7 item measure. The internal consistency reliability estimate obtained in the present study was .77, which is consistent with past research findings (Davis, 1980). The Self Monitoring Scale (Snyder & Gangestad, 1986) was used to evaluate self-monitoring. An internal consistency reliability estimate of .70 is typical in past research (Snyder & Gangestad, 1986) and it was .69 in the present sample. Private and public self-consciousness were measured by the Self-Consciousness Scale (Fenigstein, Scheier, & Buss, 1975). The public self-consciousness scale contains seven items and had an internal consistency reliability estimate of .76. The private self-consciousness scale consists of 10 items and had an internal consistency reliability estimate of .54 in the present study.

Teamwork schema similarity. Teamwork schema similarity was assessed using a measure developed for this study based on pilot research (Pape, 1997). The measure was developed using the procedure advanced by

Rentsch (1993). The measure consisted of 15 items that assessed teamwork schema. Each team member rated the 15 items for how important they were to his or her concept of teamwork. Then each team member completed the items as he or she believed his or her teammate would rate them.

Team Performance. Eight aspects of team performance were coded: problem space identification, constraint violations, misconceptions, confusions, math errors, speed, accuracy, corrections, and solution quality. Similar types of performance measures for the employed task have precedence in the research literature (e.g., McNeese, 1993; Vye, Goldman, Voss, Hmelo, & Williams, undated). Two raters coded performance independently by reviewing each team's written work and the videotape of the team. Interrater reliability was 93% agreement. The raters resolved any discrepancies in ratings through consensus.

Team processes. Team processes were coded using a coding scheme developed for this study. Two categories of team processes were assessed: team interaction processes and group processes. Team interaction processes included: depth of meaning in informing oriented communication, depth of meaning in learning oriented communication, quality of exchange, consensus evaluation communication, egocentricity versus mutuality, openness and acceptance of communication. Group processes were: group cohesion, task motivation, and team metacognition.

The total time during which the team worked was divided into quarters. Each process variable was rated for each quarter. The team received a rating for depth of meaning in informing oriented communication, depth of meaning in learning oriented communication, quality of exchange, consensus evaluation communication, egocentricity versus mutuality, group cohesion, and team metacognition for each quarter. Each team member was rated on depth of meaning in informing oriented communication, depth of meaning in learning oriented communication, consensus evaluation communication, openness and acceptance of communication, task motivation and team metacognition for each quarter.

Each variable was rated by two raters. One rater prepared a transcript of the videotape and then reviewed the videotape using the transcript to rate the team members and the team on the process measures. The other rater

used the transcripts when reviewing the videotape to make the ratings. The raters achieved interrater reliability of 77% agreement on all of the ratings. All discrepancies were resolved through consensus.

Preliminary Results

Variables

The results reported here are very preliminary. The similarity among team members on the individual difference measures was assessed using the within team standard deviations on each variable. Team level on the individual difference measures was assessed using the within team level mean for each variable. Two forms of TMSS accuracy and agreement were calculated. One form was based on absolute difference scores. Using the absolute difference score method, TMSS agreement was calculated as the average of the summed absolute difference scores between members' ratings on each teamwork schema item. TMSS accuracy scores were calculated as the average of the summed absolute differences scores between each member's ratings of the other member's schema and the other member's actual schema ratings. Stated differently, the absolute difference was calculated for each item, then summed for each team member, and the sums were averaged for the team level accuracy measure. The second form of TMSS was calculated using the square root of the sum of the squared difference scores. Both forms of TMSS produced similar results, therefore, the results from the absolute difference form are presented.

Each process variable was aggregated across quarters for a total score. An aggregated team level score was computed for those variables on which each team member was assessed. All analyses involving process variables are based on a sample size of 30. All other analyses are based on a sample size of 45.

Individual Differences as Predictors of TMSS

The similarity and level of the individual difference scores were correlated with TMSS accuracy and agreement scores. The bivariate correlations revealed no statistically significant relationships. Regression analyses were conducted using a backward stepwise strategy. These results revealed that TMSS Accuracy was predicted by team member similarity on trust and on public self-consciousness ($\underline{R} = .37$, $\underline{p} < .05$; $\beta = .24, .28$, respectively, $\underline{p} > .05$). TMSS Agreement was predicted by team level on team experience, trust, public self-consciousness, and

private self-consciousness ($\underline{R} = .45$, $\underline{p} < .05$; $\beta = -.24, .26, .28, -.38$ respectively). Private self-consciousness and team experience had negative beta weights. These are unexpected results. Additional analyses are required to test for suppression effects.

Team Interaction Processes as Predictors of TMSS

The only team interaction process variable that predicted TMSS was the aggregated team level openness variable. This variable predicted both TMSS Accuracy and Agreement with correlations of .44 ($\underline{p} < .05$). This result was unexpected, because it was in the opposite direction of the prediction. Remember a low value indicates high TMSS. Therefore, the positive correlations suggest that low TMSS is related to a high level of openness. After reviewing this result, the process coders suggested that perhaps, the measure of openness was assessing some other variable such as acquiescence or passivity. They believed that this may be the case, because team members were given high ratings on openness and acceptance when they were willing to consider the other team members' input. A high degree of openness and acceptance was characterized by immediate discussion of another's input. The more persuasion required before input is considered or discussed, the lower the level of openness and acceptance. These behaviors might also characterize high levels of acquiescence or passivity. If this is true, then these correlations may make sense.

TMSS Accuracy and Agreement as Predictors of Team Performance

TMSS Accuracy correlated significantly with how completely teams identified the problem space, with the solution rating, and with making a constraint violation ($\underline{r} = -.33, -.30, .36$, $\underline{p} < .05$). Higher TMSS accuracy was related to better problem identification and a higher solution rating. However, it was an unexpected finding that high TMSS Accuracy was related to making more constraint violations.

Simultaneous regression equations were computed in which TMSS accuracy and agreement were entered as predictors of team performance. Significant results were obtained for a constraint violation and for problem space identification. In both cases TMSS Accuracy had a significant beta weight and TMSS Agreement did not. Furthermore, in both cases, the multiple R was .39. Although an additional hierarchical regression analysis is required, it appears that TMSS Accuracy alone predicts as well as TMSS Accuracy and Agreement combined.

Team Interaction Processes as Predictors of Team Performance

Correlational analyses have been conducted to provide preliminary results regarding the nature of the relationship between team interaction processes and team performance. Problem space identification was correlated significantly with aggregated team level openness, team depth of informing oriented communication, and team learning oriented communication, and team cohesion (r = -.39, .42, .36, -.51, respectively, $p < .05$). Team cohesion and team level openness were related negatively to team performance. Again, this pattern of results may indicate that the openness measure really assessed acquiescence. It might also suggest that highly cohesive teams suffered from a bit of groupthink causing them to ignore some information.

Aggregated team openness, team cohesion, and quality of exchange were correlated significantly with constraint violations (r = .41, .38, .47, respectively, $p < .05$). All of these correlations were in the direction opposite of that predicted. Again, the acquiescence and groupthink hypotheses may explain these results.

Egocentricity versus mutuality and team depth of meaning in informing oriented communication predicted corrections (r = -.37, .40, respectively, $p < .05$). Egocentricity also predicted misconceptions ($r = -.47$, $p < .05$). Solution rating was predicted significantly by team depth of information processing ($r = .38$, $p < .05$).

Discussion Based on Preliminary Results

We would like to remind the reader that these results are preliminary for several reasons. First, the analyses involving the process variables were based on a sample size of 30. Data from all 45 teams will be analyzed in the near future. Second, additional analyses are required to further explore the nature of the predicted relationships and the nature of the unexpected results. Some of the analyses to be conducted were mentioned in the results section. Third, our dataset contains variables that may serve as control variables (e.g., time to solution, demographic characteristics). These variables have not yet been analyzed. Therefore, due to the preliminary nature of the data analysis, our discussion of these results is presented tentatively and cautiously.

In this study we tested a model of team member schema similarity (TMSS) that included individual difference and team process variables as predictors of TMSS, and TMSS and team process variables as predictors of team performance. The results revealed that the individual difference variables predicted TMSS in combination

only. In addition, similarity of individual differences predicted TMSS Accuracy and level of individual differences predicted Agreement. These results reflect the complex nature of personality. Also, we suggest that future research include the Armstrong Laboratory Aviation Personality Survey (ALAPS; Retzlaff, King, McGlohn, Callister, 1996) as a measure of individual differences related to teams and collaborative work. The variables on the ALAPS may be related more strongly to TMSS than the individual difference variables assessed in the present study.

The process variable related to openness or acquiescence predicted TMSS. This relationship requires additional study. However, we hesitate to interpret any of the process correlations in depth due to the preliminary sample size.

In past research, team effectiveness has often been assessed by the teams rating their own performance. This practice is common (Tannenbaum, Beard, & Salas, 1992). However, Tannenbaum et al. (1992) advocate the use of objective measures. In the present study, we employed objective performance measures. TMSS was not a strong predictor of these performance measures. Interestingly, included on the background survey was a single item on which team members rated their satisfaction with the teamwork that had occurred on their team. This item correlated significantly with TMSS accuracy and agreement (r = -.37, -.35, respectively).

Although based on only part of the dataset, it appears that team processes have weak relationships to TMSS. However, team processes do predict team performance. The nature of these relationships requires additional analyses.

This study has revealed that TMSS Accuracy and Agreement predict team performance differently and are related differentially to individual differences. TMSS Accuracy deserves additional investigation.

Future Research

The next step in this line of research is to test the TMSS model within the context of an ecologically valid synthetic task environment, such as the VIPER/VECTOR tasks (McNeese, personal communication) that were developed in Armstrong Laboratory, or such as the TRAP task (Wilson, McNeese, Brown, & Wellens, 1987) modified to realistically portray the UCAV or the command-post environment. This research will contribute to the

Collaborative Systems Technology Lab at The Fitts Human Factors Engineering Laboratory at the Collaborative Systems Technology Branch of Armstrong Labs at Wright Patterson AFB.

Baker and Salas (1992) suggest that teamwork may evolve, therefore researchers need to develop teamwork measures that are generalizable to teams in many settings and different team tasks. Our efforts to develop team interaction process measures are in line with this objective. Our measures will be useful in understanding the development of teamwork behaviors and the development of TMSS. We wish to pursue future research in which we will use these measures to study the development of teamwork and TMSS.

Future Applications

The theoretical constructs and the measures that we have developed as part of the Air Force Office of Scientific Research Summer Faculty Research Program applied within the ecologically valid synthetic task environment will provide the foundation for future research on Joint Cognitive Systems. As we acquire knowledge of team member schema similarity and its relation to shared responsibility and team performance, we can pursue research relating these variables to collaborative systems technologies such as datawalls or knowledge rooms. Understanding how team members interact will facilitate the design and development of human-machine communication (Wellens & McNeese, 1987).

We know that the *New World Vistas Air and Space Power for the 21st Century* publication indicates that collaborative computing is a high priority. Collaborative computing may take the form of advanced "groupware," which may be used to facilitate group interaction and decision making among co-located or distributed group members, and human-machine interaction. Clearly, understanding the effects of the joint cognitive system designs on TMSS will enhance the utility of these technologies. MUDs, or Multiple User Dimensions, are described as "....in effect social virtual realities"...."to manage the details of the social interaction... (NWV, p. 100)." The New World Vistas report goes so far as to suggest that a "virtual Pentagon (p. 101)" may become a reality. These innovations require extensive knowledge of socio-cognitive processes and their relationships to technology.

Footnotes

1. Perceiver refers to the individual who is asked to rate another person. The person being rated is referred to as the target.

References

Baker, D. P., & Salas, E. (1992). Principles for measuring teamwork skills. Human Factors, 34(4), 469-475.

Bettenhausen, K. L. & Murnighan, J. K. (1991). The development of an intragroup norm and the effects of interpersonal and structural challenges. Administrative Science Quarterly, 36, 20-35.

Brown, C. E., Whitaker, R. D., Selvaraj, J. A., & McNeese, M. D. (1995). The development of trace: An integrative bargaining paradigm for investigating multidisciplinary design tradeoffs. Air Force Material Command, Wright-Patterson Air Force Base, Ohio. (Al/CF-TR-1995-0073).

Cannon-Bowers, J. & Salas, E. (1990, April). Cognitive psychology and team training: Shared mental models in complex systems. Paper presented to the Meeting of the Society for Industrial/Organizational Psychology, Miami Beach, Florida.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), Individual and group decision making (pp. 221-246). Hillsdale: Lawrence Erblaum Associates.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS: Catalog of Selected Documents in Psychology, 10, 85.

Duffy, L. (1993). Team decision making and technology. In N. J. Castellan (Ed.), Individual and group decision making (pp. 247-265). Hillsdale: Lawrence Erblaum Associates.

Dyer, J. L. (1984). Team research and team training: A state of the art review. In F. A. Muckler (Ed.), Human factors review (pp. 285-323). Santa Monica: Human Factors Society.

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. Journal of Consulting and Clinical Psychology, 43(4), 522-527.

Fischer, U. M. (1996). Methods for analyzing group problem solving decision making. United States Army Research Institute for the Behavioral and Social Sciences. (ARI Research Note 96-64).

Gersick, C. (1989). Marking time: Predictable transitions in task groups. Academy of Management Journal, 32, 274-309.

Hakel, M. D., Weil, E. K., & Hakel, L. (1988). The analysis and clustering of navy ratings based on social interaction characteristics: A literature review and conceptual model. Naval Personnel Research and Development Center, San Diego, California. (TR-88-46).

Kimble, C. E., & McNeese, M. D. (1987). Emergent leadership and team effectiveness on a team resource allocation task. Harry G. Armstrong Aerospace Medical Research Lab, Wright-Patterson Air Force Base, Ohio. (AD-A192-105).

McNeese, M. D. (1993). Putting knowledge to use: The acquisition and transfer of knowledge in situated problem solving environments. Dayton, OH: Wright-Patterson Air Force Base, Armstrong Laboratory. (NTIS No. AL/CF-TR-1993-0052).

McNeese, M. D., Zaff, B. S., Citera, M., Brown, C. E., & Whitaker, R. (1995). AKADAM: Eliciting user knowledge to support participatory ergonomics. The International Journal of Industrial Ergonomics, 15(5), 345-363.

New World Vistas: Air and Space Power for the 21st Century. Information Technology Volume. Department of the Air Force. USAF Scientific Advisory Board.

Nosek, J. T., & McNeese, M. D. (1997). Issues for knowledge management from experiences in supporting group knowledge elicitation and creation in ill-defined, emerging situations. Paper presented to the AAAI Spring Symposia: Knowledge Management, Standford University, March.

Pape, L. J. (1997). The effect of personality characteristics on team member schema similarity. Unpublished master's thesis proposal. Wright State University, Dayton, Ohio.

Poole, M. S., Gray, B. & Gioia, D. A. (1990). Organizational script development through interactive accommodation. Group and Organization Studies, 15(2), 212-232.

Poole, M. S. & McPhee, R. D. (1983). A structural analysis of organizational climate. In L. L. Putnam & M. E. Pacanowsky (Eds.), Communication and organizations: An interpretive approach (pp. 195-219). Beverly Hills, CA: Sage Publications.

Rentsch, J. R. (1993, August). Predicting team effectiveness from teamwork schema similarity. Paper presented at the Academy of Management Meetings, Atlanta, Georgia.

Rentsch, J. R., & Hall, R. J. (1994). Members of great teams think alike: A model of team effectiveness and schema similarity among team members. In M. M. Beyerlein & D. A. Johnson (Eds.), Advances in interdisciplinary studies of work teams. Vol. 1. Series on self-managed work teams (pp. 223-262). Greenwich, CT: JAI Press.

Rentsch, J. R., Pape, L. P., & Brickman, N. M. (1997). Team member schema similarity: Exploring individual differences predictors. Unpublished manuscript. Wright State University.

Retzlaff, P. D. King, R. E., McGlohn, S. E., & Callister, J. D. (1996). The development of the Armstrong Laboratory Aviation Personality Survey (ALAPS). Air Force Material Command, Brooks Air Force Base, Texas. (AL/AO-TR-1996-0108).

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. Journal of Personality, 35, 651-665.

Schneider, B. (1987). People make the place. Personnel Psychology, 40, 437-453.

Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. Journal of Personality and Social Psychology, 51(1), 125-139.

Tannenbaum, S.I., Beard, R.L., & Salas, E. (1992). Team building and its influence on team effectiveness: an examination of conceptual and empirical developments. In K. Kelley (Ed.), Issues, theory, and research in industrial/organizational psychology (pp. 117-153). North-Holland: Elsvier.

Tower, S. L., & Elliott, L. R. (1996). The role of communication efficiency in teams with distributed expertise: Application of the multi-level theory. Proceedings of the International Association of Management Conference, Toronto, Canada.

Vye, N. J., Goldman, S. R., Voss, J. F., Hmelo, C., Williams, S., & Cognition and Technology Group at Vanderbilt. (undated). Complex Mathematical Problem Solving by Individuals and Dyads. Unpublished manuscript. Vanderbilt University, Nashville, Tennessee.

Walsh, J. P., Henderson, C. M, & Deighton, J. (1988). Negotiated belief structures and decision performance: An empirical investigation. Organizational Behavior and Human Decision Processes, 42, 194-216.

Wellens, A. R., & McNeese, M. D. (1987). A research agenda for the social psychology of intelligent machines. IEEE Systems, Man, and Cybernetics.

Wells, M. J., & Hoffman, H. G. (1996). The virtual crewmember. Paper presented to the Human Factors and Ergonomics Society, Philadelphia, September 2-6.

Wilson, D. L., McNeese, M. D., Brown, C. E., & Wellens, A. R. (1987). Utility of shared versus isolated work setting for dynamic team decision-making. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. (AD-A192-434).

Young, M. F., & McNeese, M. D. (1995). A situated cognition approach to problem solving. In Hancock, P., Flach, J., Caird, J., & Vincente, K. (Eds.), Local applications of the ecological approach to human-machine systems. Hillsdale: Lawrence Erblaum Associates.

Figure 1: Team Member Schema Similarity Model (Rentsch & Hall, 1994)

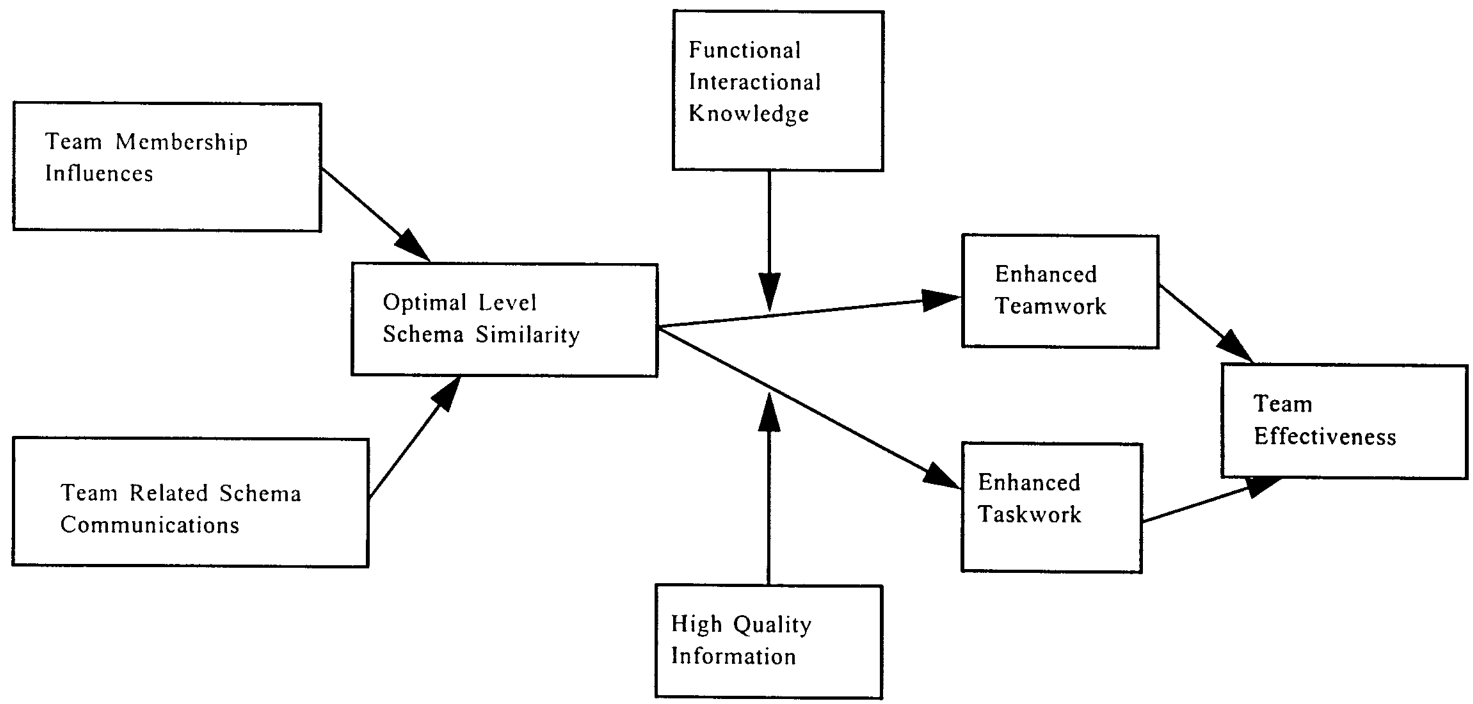
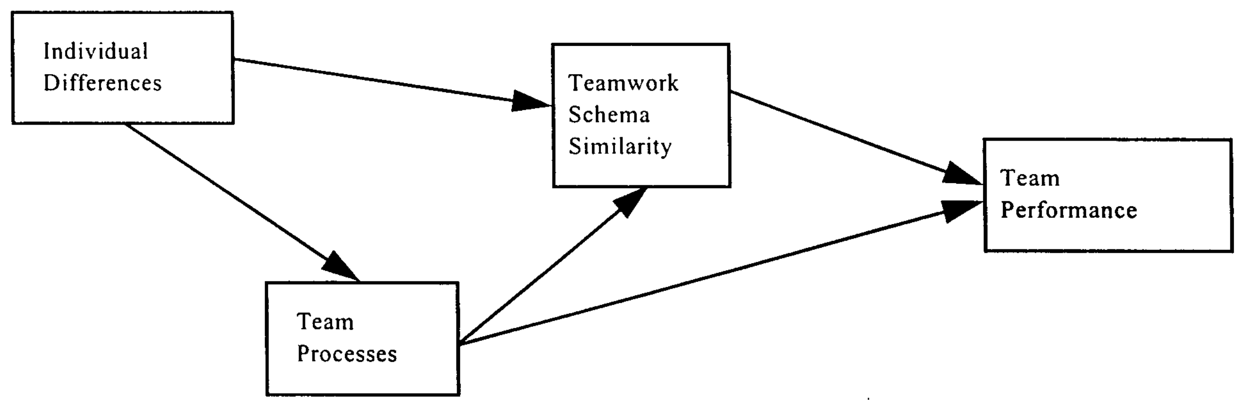


Figure 2: Model Modifications



PERCEPTION OF VELOCITY AS A FUNCTION OF THE OCULOMOTOR STATE OF THE EYES

Bradley E. Collie
Graduate Student
Department of Aeronautical Technology

Arizona State University
School of Technology
6001 South Power Road - SIM Building 425
Mesa, Arizona 85206

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

PERCEPTION OF VELOCITY AS A FUNCTION OF THE OCULOMOTOR STATE OF THE EYES

Bradley E. Collie
Graduate Assistant
Department of Aeronautical Technology
Arizona State University

ABSTRACT

Two studies investigated velocity perception as a function of the oculomotor state of the eyes (vergence and accommodation). In the first experiment, images of moving simulator terrain were viewed on collimated and real (non-collimated) displays. The collimated display induced little to no vergence and the real display, at close viewing distances, induced a large amount of vergence. When the images in the collimated and the real displays were set at equal velocities, the collimated image appeared faster than the real image. Thus the velocity of the real image had to be increased in order for the two images to be perceived as identical in velocity. A pilot study is currently underway to identify any problems associated with this study.

In the second experiment, a moving texture pattern was viewed on flat surface screens placed at varying distances from the observer. The primary focus of this experiment was to examine how changes in the oculomotor state of the eyes effect the perception of velocity. The far display (viewed at 5.5 m from the observer) induced less ocular vergence and appeared faster than either of the near targets (viewed at 1.2 m or .5 m from the observer). Thus the velocity of both near targets had to be increased in order for the images to be perceived as identical in velocity. Data collection has been completed for the multiple-linear and multiple-radial motion, and is currently being collected for the multiple-circular motion.

PERCEPTION OF VELOCITY AS A FUNCTION OF THE OCULOMOTOR STATE OF THE EYES

Bradley E. Collie

EXPERIMENT ONE

INTRODUCTION

The effects of optical vergence and accomodation are currently being investigated in an effort to further understand the problems associated with the perception of simulator imagery.

A study by Pierce, Geri and Hitt (in press) examined how collimated and non-collimated display types effect the perception of size and distance of an object. High-resolution F-15 aircraft that remained stationary to the observer were superimposed on moving simulator background imagery. Reducing or enlarging the aircraft image proportionate to the estimated viewing distance represented distance of the aircraft. The images were viewed from the side, as in formation flight, or from behind and above as during a gun pass. The results indicated, for simulator imagery, that both perceived size and perceived distance are significantly reduced when imagery is displayed at viewing distances of less than one meter as compared to when it is displayed at or near optical infinity.

The current study expands on the size and distance estimation properties observed in the Pierce, et al. (in press) study. With the known differences of size and distance perception associated with each type of display, we examined how these differences effect the perception of velocity. Using simulator background we compared the differences in perceived velocity of moving terrain in a flow motion (as you would see the terrain moving if you were looking out the forward window in an aircraft in straight and level flight) on collimated and real displays.

There were two independent variables in this study. The first was the type of display (collimated and non-collimated). The second was velocity. For each trial, a collimated image

was shown paired with a real image; one image was set at a standard velocity and the other velocity was variable. The variable velocity was initially predetermined, however subsequent velocities were a function of the previous response made in that staircase. For each session there were two staircases. One staircase set the collimated image at the standard velocity and the other staircase set the real image at the standard velocity. Order of presentation (collimated and real) was determined randomly for each trial. After seven reversals, that staircase ended and the last six reversals were averaged.

METHOD

Subjects. Although the subject pool has not been completed, this study will consist of approximately 4 observers with corrected to normal vision.

Apparatus/Materials. The displays were aligned at a 90-degree angle from each other and were viewed through a beamsplitter. The collimated display consisted of a Wide-Angle Collimated (WAC) window that displayed background imagery from a Barco Graphics 808 CRT projector with close focus lens.

Insert Figure 1 about here

A headrest was placed 18.75 cm in front of the beamsplitter, thus allowing the observer's head to remain stationary while viewing the two displays through the beamsplitter. The collimated display was placed directly in front of the observer and measured 106.25 cm from the beamsplitter (measured center of WAC window semi-silvered glass to center of beamsplitter).

Insert Figure 2 about here

The viewing distance from the observer's eye was 125 cm for the collimated display and 83.75 cm for the non-collimated display. The real (non-collimated) display was placed to the right of the observer and measured 65 cm from the beamsplitter (measured center of the screen to the center of the beamsplitter). The real display consisted of a plastic, translucent screen (100 cm x 75 cm) which displayed background imagery from a Barco Graphics 808 CRT projector with close focus lens.

Insert Figure 3 about here

A 3-button mouse was used as a response box (the ball was removed and only 2 of the 3 buttons were used for this study).

The experiment was set-up in a large computer room and a curtain wall had to be erected to completely enclose the testing area to eliminate extraneous light from entering. The testing area was located in a corner of the room, so the two existing walls were used and only 2 curtain walls had to be erected. The two walls consisted of aluminum pipe, which we drilled holes into and attached Velcro via a rivet gun. Black curtains with Velcro sewn to the outside edges were attached to the poles. Each wall measured 10'9"h X 18'w.

Procedure. The observer was given instructions about the procedures of this study and how to make an appropriate response using the response box. The observers were seated in the apparatus in a dark room and placed their head in a chinrest. Upon starting the session, the observer viewed a sequence of two displays (one from the collimated display and one from

the non-collimated display) of simulated moving terrain background. The observer viewed the first display for 5 seconds, then the second display was viewed for 5 seconds. The observer then made a response as to whether the second display was faster or slower than the first. Responses were made via a 3-button mouse, where the right button indicated "faster" and the left button indicated "slower". The session was completed and automatically terminated by the computer when the observer had made 7 reversals on each of the two staircases.

EXPERIMENT TWO

INTRODUCTION

In the second experiment we examined how variations in viewing distance effected velocity perception of radial, linear and circular moving patterns. As the distance between an observer and the object being viewed increases, the amount of ocular vergence required to view that object decreases.

In this study, a moving texture pattern was projected at different velocities onto the backside of two flat screen surfaces. An image from the far distance condition was viewed and compared with an image from one of the two near distance conditions (depending on which near condition was being tested). Observers were asked to decide if the second image was faster or slower than the first.

There were three independent variables in this study. The first was the variable viewing distances. The standard distance was constant at 5.5 meters and the variable distance was either 1.2 meters or 0.5 meters. The second independent variable was velocity. Velocities were standardized as Low, Medium and High (velocities of 3, 6 and 9 degree's per second were used in the Multi-Radial motion and corresponding velocities of 6, 12 and 18 degree's per second were used in the Multi-Linear and Multi-Circular motion). The third independent variable was the type of motion. There were three types of motion used in this study: 1)

linear motion (dots move across the screen horizontally), 2) radial motion (the dots originate anywhere on the screen and move outward) and 3) circular motion (the dots circle at a constant distance around the fixation point in the center of the screen).

METHOD

Subjects. The study consisted of 4 male observers.

Apparatus/Materials. The two displays were aligned at a right angle to each other and were viewed through a beamsplitter so the observer could view both displays without moving his head. This configuration is similar to that seen in Figure 1, with the exception that both images were displayed on flat screen surfaces and viewing distances were different. The far display was placed to the right of the observer at a distance of 5.5 meters from the beamsplitter. The far display consisted of a Barco Graphic 800 video projector and a plastic screen (95"w x 75"h). The near displays were placed directly in front of the observer at a distances of 1.2 meters or .5 meters from the beamsplitter, depending on which condition was being tested. The near display consisted of a Barco Graphics 800 projector and a plastic screen (48"w x 36"h). However, the backside of the screen was occluded with cloth to reduce the screen viewing size to 29.5"w x 24.5"h.

The observer station was inside an enclosed curtain area with 4 walls and a cover over the top to conceal the observer from extraneous light. There were openings in the curtain to allow the image from the two displays to be viewed through the beamsplitter, however, occluders on the right side of the observer prevented him from seeing the screen directly, thus eliminating any cues as to which display was being shown through the beamsplitter. A headrest was placed 18 cm in front of the beamsplitter. This allowed the observer to view both displays without moving his head. A response box with a toggle switch was used by the observer to indicate whether the second stimuli was faster (push toggle forward) than the first or slower (pull toggle back).

An eye-tracker was used to measure ocular vergence.

Procedure. The observers were familiarized with the apparatus and the response box. Then the lights were turned off and the observer was allowed to dark adapt for a period of approximately 10 minutes. Ocular vergence was measured using an eye-tracker. Upon starting the session, the observer would view a sequence of two displays of stimuli, one from the far display and one from the near. The observer viewed the first display for 5 seconds. After, the second display was viewed for 5 seconds. The observer then made a response as to whether the second display was faster or slower than the first. Responses were made via a toggle switch, where pushing the toggle forward indicated the second stimuli was "faster" than the first, and pulling the toggle backward indicated "slower". The computer paused after each pairing of stimuli, giving the observer time to make a response. After the observer had made a response, the computer continued with the next pair of stimuli. The session was completed and automatically terminated by the computer when the observer made seven reversals in each of the two staircases. The last six reversals were averaged.

GENERAL DISCUSSION

The raw data generally conform for these two experiments, provided all other variables remain the same, as vergence increases, perception of velocity decreases. It still remains to be seen whether or not the differences between the variables in this study are statistically significant.

The methods used to manipulate vergence varied between the two experiments (Experiment One used collimated and real displays, and Experiment Two used viewing distance). However, when images were set at equal velocities and were viewed on the display requiring less vergence (as in the collimated display in Experiment One or the far display in Experiment Two) the image appeared faster than when viewed on the display requiring greater vergence (as in the real display in Experiment One and either of the two near displays in Experiment Two). Thus, the velocity of the image requiring greater vergence must be

increased in order for it to be perceived as the same velocity as the image requiring less vergence.

SUMMARY

My responsibilities during my summer program were to coordinate with the appropriate engineering support elements to ensure that each experimental apparatus was calibrated, aligned, and that the software was set up according to the specific needs of the study. Furthermore, I was responsible for gathering and testing subjects, data collection, analysis and presentation, as well as various administrative duties.

My specific accomplishments were: I assembled two curtain walls out of aluminum pipe, Velcro and black curtains for use in the WAC study; I was responsible for ensuring that the engineering support element properly aligned the images in the WAC and Real displays and that the Barco projectors were properly color matched for use with the AVTS computer system; I worked with our computer programmer to set up the software according to the needs of the study (such as constructing the staircases and step sizes, randomization of the standard velocity and the order of display, collection of essential data, and many others); I began a pilot study in Experiment One, identified potential problems and corrected them; Upon replacing the old AVTS computer system with the new SE 2000 in Experiment One, I was responsible for coordinating with the appropriate engineering support elements to align and color match the two displays, and to assist our programmer in the reconstruction of the computer program to meet the specific needs of the study; I began a pilot study with the new computer system to identify and correct any problems associated with this study; I gathered and pre-screened subjects to be used in our subject pool and scheduled them to come in when necessary for testing; I collected data in Experiment Two for the linear and radial motions at 0.5 and 1.2 distances and began testing for circular motion; I analyzed data collected from current and previous studies; I created and modified graphs and tables for publication as needed; I performed various administrative duties such as typing various reports, scheduling time for use on the SE 2000 computer system, submitting timesheets for paying the subjects,

and so on; and finally, when individuals or groups came in for a tour of the laboratory, I assisted in giving demonstrations of the study described in Experiment One or Two.

REFERENCES

Pierce, B., Geri, G., & Hitt, J. (in press). The Perceived Size and Distance of Collimated and Real Flight Simulator Imagery (AL/HR-TR-xxx). Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

CAPTIONS

Figure 1. Apparatus placement for Experiment One.

Figure 2. WAC window configuration used as the Collimated display in Experiment One.

Figure 3. Flat screen display configuration used as the Real display in Experiment One. A similar configuration was used for both displays in Experiment Two.

Figure 1.

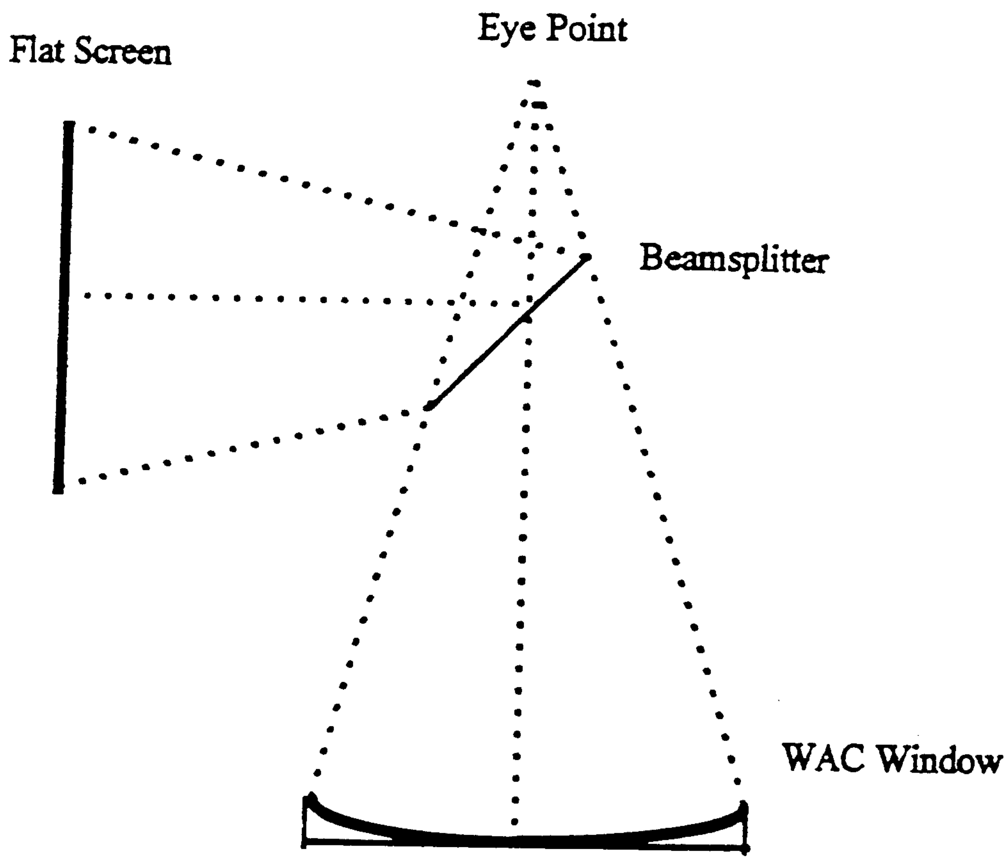


Figure 2.

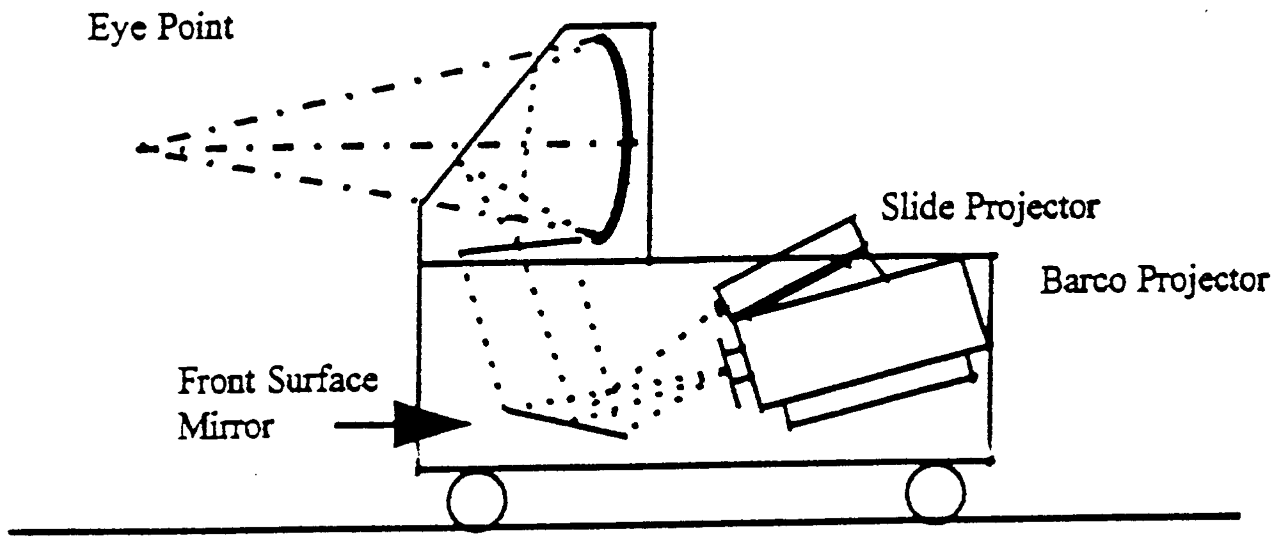
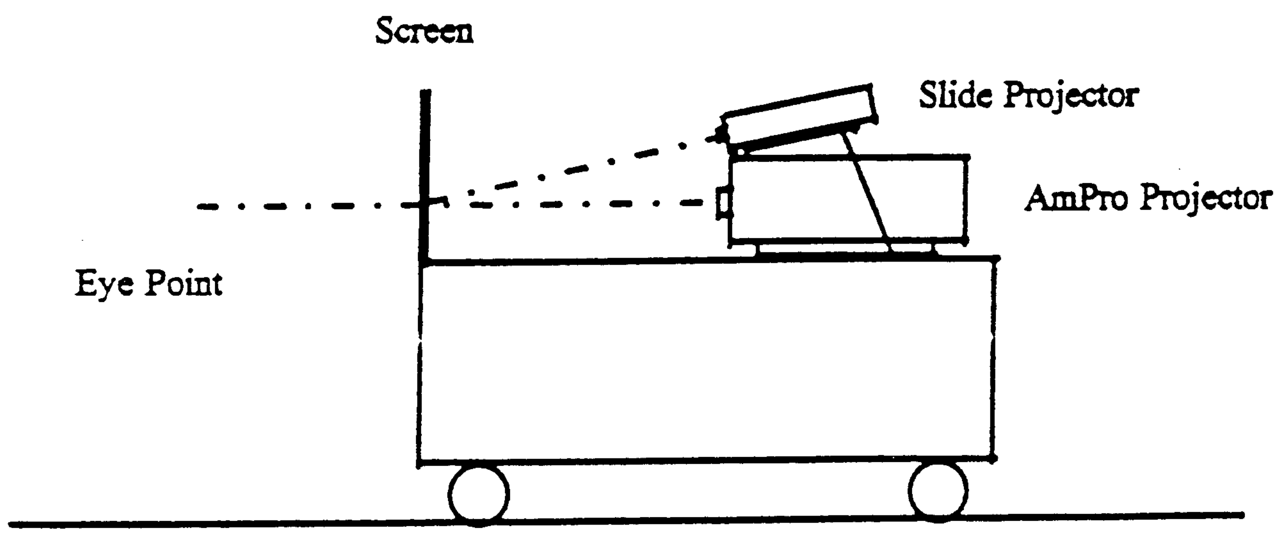


Figure 3.



INVESTIGATION AND VALIDATION OF SUBMAXIMAL CYCLE ERGOMETRY PROTOCOLS USED TOASSESS THE AEROBIC CAPACITY OF THE USAF POPULATION

Gerald W. DeWolfe
Graduate Student
Exercise Physiologist

Department of Kinesiology
The University of Texas at Austin
222 Bellmont Hall
Austin, TX 78712

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

Table of Contents

Investigation and Validation of Submaximal Cycle Ergometry Protocols Used to Assess the Aerobic Capacity of the USAF Population

Gerald W. DeWolfe
Exercise Physiologist
Department of Kinesiology
The University of Texas at Austin

Abstract

The United States Air Force (USAF) uses a submaximal cycle ergometry protocol to assess the aerobic fitness level ($\dot{V}O_2$ max) of all members. Weaknesses of the current assessment require either modifications to or replacement of the current protocol. The purposes of this study were to evaluate (1) the Banister-Legge (B-L) protocol (men only), (2) a modified USAF assessment with (USAF+CF) and without a correction factor (USAF), and (3) a new steady state (SS) protocol using the current Air Force equation (SS+CF and SS) to generate $\dot{V}O_2$ max estimates. The B-L protocol significantly underestimated the $\dot{V}O_2$ max of all male subjects. The USAF+CF and USAF significantly underestimated actual $\dot{V}O_2$ max for male subjects. The USAF+CF had a lower mean difference, SEE, and E than the USAF estimation. The USAF+CF significantly underestimated the actual $\dot{V}O_2$ max for females, but the USAF was not significantly different than actual $\dot{V}O_2$ max. The SS and SS+CF both significantly underestimated $\dot{V}O_2$ max for men and women. The mean difference, SEE, and E were all higher compared to the USAF and USAF+CF estimates.

The best $\dot{V}O_2$ max estimates for male subjects in this study were USAF+CF and the best for female subjects were the USAF estimates. Due to the differences between the USAF and SS protocol, using the USAF equation with SS test responses may not be appropriate. With a new prediction equation based upon the SS test responses, the SS test should produce more accurate $\dot{V}O_2$ max estimates. The B-L protocol is not a feasible replacement for the USAF protocol.

INTRODUCTION

Submaximal cycle ergometry has been used for over 40 years to predict maximal oxygen consumption ($\dot{V}O_2$ max) (1). $\dot{V}O_2$ max is defined as the maximum amount of oxygen a person can consume during a graded exercise test to exhaustion. This value is considered one of the best indicators of an individual's aerobic fitness level. $\dot{V}O_2$ max can be measured directly by analyzing expired gases and ventilatory volumes from a subject exercising to exhaustion, normally on a treadmill or cycle ergometer. This process is considered to be the most accurate and reliable method for determining an individual's aerobic-fitness level. However, due to cost, time, manpower, and a small risk of injury or death, $\dot{V}O_2$ max tests are not practical for administration to large numbers of people.

The United States Air Force's (USAF) mission is to defend the United States through control and exploitation of air and space (Air Force Mission Statement). To fulfill this mission each individual must possess the physical capacity, including aerobic fitness level, to complete duties that involve physical work. The USAF commanders and fitness program scientists have determined a minimum $\dot{V}O_2$ max each person should possess to be able to fulfill mission requirements. Establishing a minimum standard necessitates identification of a practical method to assess aerobic fitness in the USAF population.

The initial technique used to assess aerobic fitness level in the USAF was a 12 minute run for distance test developed by Dr. Kenneth Cooper (3). Dr. Cooper developed an equation to predict $\dot{V}O_2$ max from distance covered in 12 minutes. Distance covered during the assessment was stratified from the lowest (Category 1) to the highest (Category 5) aerobic fitness category, with category 3 being the minimum requirement to pass. Individuals completing 1.25 miles or greater within the 12 minute time limit passed the assessment. Individuals who did not meet the distance requirement received a fail.

The Cooper test suffered from too many negative components including the following: risk of injury or death due to extreme temperatures and low levels of fitness in some USAF personnel, non-standardized testing conditions among Air Force Bases (AFB), and assessment standards not adjusted for sex or age. Due to the limitations of the 12 minute run/walk assessment the USAF sought to develop a safer, more objective test to determine the fitness level of their population. It was decided a submaximal test to estimate $\dot{V}O_2$ max was the most feasible method of assessment and in 1992, a submaximal cycle ergometry test was adopted to assess the aerobic fitness level of the USAF population.

The USAF adopted a submaximal cycle ergometry protocol modeled after the original Astrand and Ryhming protocol (1), which in essence is a 6-minute steady state protocol designed to determine a heart rate (HR) response to a given workrate (WR). This protocol consists of a

warmup phase, WR progression phase, steady-state period, and cool-down period. The HR and WR at the end of the steady-state period are then used to generate a $\dot{V}O_2$ max estimate using an algorithm developed by Dr. Loren Myrhe (10).

The USAF has used the current cycle ergometry protocol for approximately five years. During this time, shortcomings of this protocol have been identified which include: a higher than desired standard error of estimation (SEE) when predicting $\dot{V}O_2$ max and the production of an excessive number of invalid assessments (11). Invalid assessments occur either due to a HR response outside of test parameters or a technical error. The majority of invalid assessments are due to the HR response of the subject being: 1) below set limits (125 bpm); 2) above set limits (> 85% predicted HRmax); or 3) too variable at the end of the six minute steady-state period ($> \pm 3$ bpm over the last 2 minutes) (4). Technical error may occur as a result of computer or tester error during administration of the assessment.

Due to the identified shortcomings of the USAF assessment, alternative methods to improve the accuracy and validity of the $\dot{V}O_2$ max estimate are needed. Three new protocols, Banister-Legge (B-L), Ramp (R) and Steady State (SS), have been identified as possible replacements for the current submaximal assessment. The B-L protocol estimates $\dot{V}O_2$ max using the ΔHR (difference between the ending submaximal HR and a baseline HR) and ending submaximal WR during a standardized submaximal protocol. The R protocol is designed to use the HR-WR response to a ramp WR progression. The ramp progression occurs in 5 W increments every 20 seconds to the point of 85% predicted HRmax. The SS protocol also uses a ramp WR progression to determine the proper WR for a six minute steady state period. The last two submaximal protocols are designed to operate in a totally automated assessment procedure.

The purposes of this study were: 1) directly measure $\dot{V}O_2$ max during both a treadmill and a cycle ergometer maximal test; 2) determine the accuracy of estimating $\dot{V}O_2$ max using the proposed submaximal cycle ergometry assessments; and 3) evaluate a modified version of the current USAF submaximal cycle ergometry protocol.

METHODS

Subjects. Forty-two healthy volunteer subjects (21 males, 21 females) 18-30 years of age participated in this study. The study was approved by the Institutional Review Board at The University of Texas at Austin. Subjects were recruited from Kinesiology classes and by flyers distributed across the UT campus. All subjects provided written informed consent and completed a medical screening questionnaire prior to participating in this study. Subjects were instructed to report to the lab having abstained from food, nicotine, and caffeine for at least three hours and from strenuous activity and alcohol for at least 12 hours prior to testing.

Experimental Design. Each subject completed five exercise tests over a three to four week period. A maximal treadmill test was completed during the initial visit. Over the next four visits, subjects completed three submaximal and one submax-to-max cycle ergometer test in random order. Metabolic measurements were made continuously during all tests using the SensorMedics 2900 Metabolic Measurement Cart (MMC). Pre-test calibration and post-test verification of calibration for the gas analyzers and the volume flow meter were completed for all exercise tests to maintain the accuracy of measurement.

Procedures. Upon reporting to the lab each day, subjects were weighed to the nearest 0.1 lb. using a balance scale, and fitted with a Polar HR Monitor™. All subjects filled out a daily questionnaire relating to food, nicotine, and caffeine intake within the previous three hours, physical activity and medications taken in the previous 12 hours, hours of sleep during the previous night, and a general rating of "how do you feel". Height was measured to the nearest 0.1 cm using a stadiometer during the initial visit only. Prior to each cycle ergometer test, resting HR was measured for two minutes while subjects were seated on the cycle ergometer. During the initial visit resting heart rate was measured while subjects were seated in a chair.

Maximal Treadmill Test (TM). Initially, subjects completed all paper work required to participate and receive compensation for participation in this study. After measuring resting HR, the proper seat height necessary to complete the cycle ergometer tests was set. The seat height was adjusted by having the subject sit on the bike with the heel of their foot on the pedal and leg fully extended. The subject was then instructed to place their foot inside the toe clips of the pedal and cycle for 3 to 4 revolutions. Seat height was accepted when the subject's leg was bent approximately 5 degrees at bottom of crank. This seat height was recorded and used for each subsequent cycle test.

Prior to the maximal test, subjects were allowed a short warmup period to familiarize themselves with walking on the treadmill. The initial speed for the treadmill test was set at either 3

or 4 miles per hour (mph) for females or 4 or 5 mph for males. Every two minutes the speed was increased 1 mph until the end of the two minute stage in which a respiratory exchange ratio (RER) of at least 0.95 was observed. At this point, the treadmill grade was increased 2% every minute until volitional exhaustion. At the point of volitional fatigue, the test was stopped and the treadmill speed was reduced to 2 mph to allow the subject a three minute cool-down period. The HR response during the test was recorded during the final 5 seconds of each minute. HRmax was the highest HR observed during the test. $\dot{V}O_2$ was measured continuously and reported as a rolling average of three 20 second measurements. $\dot{V}O_2$ max was the highest one minute value observed during the test using the rolling average of three 20 second measurements.

Following the treadmill test, each subject completed a familiarization trial on the cycle ergometer before leaving the lab. Subjects cycled for four minutes with no resistance (unloaded cycling) and for four minutes at 30 W for females or 50 W for males.

Cycle Submax-to-Max Test (B-L). A cycle submax-to-max test was completed on a SensorMedics 800-S Ergo-Metric cycle ergometer. The submaximal portion of this test was designed to reproduce a protocol developed by Banister and Legge (9). The protocol set forth by Banister and Legge was the following: five minutes of unloaded cycling at 90 rpm followed by WR increases of 50 watts (W) every three minutes (e.g. 50, 100, 150, 200, and 250). The Banister-Legge (B-L) protocol was designed for males aged 20-29 years of age. An alternate protocol was designed for the female subjects of this study. The female protocol was the following: five minutes of unloaded cycling at 90 rpm followed by WR increases of 25 W every three minutes (e.g. 25, 50, 75, 100, 125, and 150). For both males and females, the submaximal WR progression was maintained until the end of a three minute stage in which a steady-state HR was achieved at or near 85% predicted HRmax (HRmax was based upon the formula: HRmax = 220 - age). If it was determined the subject's HR response would not achieve steady-state before exceeding the 85% predicted HRmax level, a HR slightly lower than 85% predicted was accepted as the end point of the submax test. HR responses were also accepted at the end of the submax test if they were slightly above 85% predicted HRmax.

At the completion of the submaximal portion of the test, the WR was increased by 25 W every two minutes until volitional fatigue of the subject. Subjects completed three minutes of unloaded cycling after finishing the cycle test. The HR response during the test was recorded during the final 5 seconds of each minute. HRmax was the highest HR observed during the test. $\dot{V}O_2$ was measured continuously and reported as a rolling average of three 20 second measurements. $\dot{V}O_2$ max was the highest one minute value observed during the test using the rolling average of three 20 second measurements. WR max was the highest WR maintained for at least one minute at the end of the test.

After testing the first 21 subjects, an additional thirteen subjects were recruited to provide a more comprehensive validation of the B-L protocol. Therefore, a total of 34 male subjects completed the B-L protocol. Of these 34 subjects, only 31 subjects had a steady-state HR response at the end of the B-L test. Since $\dot{V}O_2$ max estimations can only be generated from steady state HR responses, $\dot{V}O_2$ max estimations were reported for a total of 31 subjects.

Ramp Protocol (R). The R protocol began with three minutes of unloaded cycling at 75 rpm. At the end of the unloaded cycling, the WR was set at 25 or 50 W for females and males, respectively. The WR then progressed 5 W every 20 seconds until the end of the 20 second period in which 85% predicted HRmax was observed. The WR was then reduced to 20 W for a three minute recovery period. HR was recorded at the end each 20 second period during the ramp progression and the end of each minute during unloaded cycling and the recovery period. $\dot{V}O_2$ was reported every 20 seconds during the ramp progression. During the unloaded cycling and recovery periods, the $\dot{V}O_2$ values were reported as the one-minute average of three 20 second measurements. Modeling techniques will be used to develop equations to estimate $\dot{V}O_2$ max from data collected during the R protocol. Analysis of the data has not been completed as of this writing, so no data from the R protocol will be presented in this paper.

Steady-State Protocol (SS) The SS protocol began with three minutes of unloaded cycling at 75 rpm. After the unloaded cycling period, the WR was set at 25 and 50 W for females and males, respectively. The WR then progressed 5 W every 20 seconds until the end of the 20 second period in which 70% predicted HRmax was observed. The WR which elicited 70% HRmax was then maintained for a total of 6 minutes. At the conclusion of the six minute steady-state period, subjects completed a three minute recovery period at 20 W. HR was recorded at the end of each 20 second period during the ramp progression and the end of each minute during unloaded cycling, six-minute steady-state period, and the recovery period. $\dot{V}O_2$ was reported every 20 seconds during the ramp progression and reported as the one minute average of three 20 second measurements during unloaded cycling, the six-minute steady state period, and the recovery period.

Air Force Submaximal Cycle Ergometry Assessment (USAF). The USAF assessment is a branching WR protocol based upon the HR responses of the subject being tested. The USAF assessment is designed to estimate $\dot{V}O_2$ max using the HR response between 125 beats per minute (bpm) and 85% predicted HRmax. The current USAF assessment consists of a two minute warmup period at a work rate of 25 W for females or 50 W for males, a three minute WR progression phase in which the HR response is evaluated every minute to determine necessary WR progressions, and a six minute steady-state period. Subjects complete this assessment while pedaling at 50 rpm.

A submaximal cycle ergometry study of the USAF assessment conducted by Flatten et al (5) indicated three 2-minute WR progression stages were superior to three 1-minute WR progression stages in producing valid test results. Also, the original Astrand-Ryhming protocol (1) used a $\pm$ 5 bpm difference as a steady state HR response in the last two minutes of the test. Therefore, we evaluated a new submaximal cycle ergometry protocol which contained three 2-minute WR progression stages and which allowed a $\pm$ 5 bpm difference in the last two minutes of the steady state period to be considered a valid test. A three minute recovery period cycling at 20 W was also added to the assessment.

USAF and SS $\dot{V}O_2$max estimations. $\dot{V}O_2$ max estimations for the USAF protocol are generated using a prediction equation which is made up of three separate factors. The Xa factor is designed to estimate the $\dot{V}O_2$ of the subject at the ending WR of the six minute steady state period, the Xb factor is designed to estimate the maximal HR of the subject, and the Xc factor adjusts for the age of the subject. In 1995, a cross-validation study of the USAF assessment was completed by the University of Florida (11). They determined the prediction equation overestimated higher fit females and underestimated lower fit males. Therefore, a correction factor (CF) was instituted to reduce female $\dot{V}O_2$ max estimates and increase the $\dot{V}O_2$ max estimates for males. The final $\dot{V}O_2$ max estimate is generated by multiplying the Xa, Xb, and Xc factors, dividing by the body weight in kilograms, and adding the CF. The USAF Fitness Program has determined the CF is no longer necessary for male subjects, so the CF will be eliminated for male subjects only beginning January 1998. Therefore, the results of both the USAF and USAF+CF $\dot{V}O_2$ max estimates were calculated.

The $\dot{V}O_2$ max estimates for the SS test were generated using the USAF+CF and USAF equations. To identify the results specific to the SS test, the results will be reported as SS+CF and SS $\dot{V}O_2$ max estimates.

Banister-Legge $\dot{V}O_2$ max estimations. $\dot{V}O_2$ max estimations of the male subjects were generated for the B-L test using tables provided by Banister and Legge (9). The B-L nomogram uses separate equations to estimate the $\dot{V}O_2$ max of trained and untrained subjects. The original study by Banister and Legge provided two methods to identify a subject as either trained or untrained: the slope of the ΔHR-$\dot{V}O_2$ and the ΔHR-%$\dot{V}O_2$ max relationships during the submax test. For the purpose of data analyses, the ΔHR-$\dot{V}O_2$ estimations and the ΔHR-%$\dot{V}O_2$ max estimations have been identified as BL#1 and BL#2, respectively. $\dot{V}O_2$ max estimations using the trained equation for all subjects were also completed (BL#3).

The B-L tables provide a $\dot{V}O_2$ max estimation in liters of oxygen consumed per minute for each ΔHR at 50, 100, 150, 200, and 250 W. The ΔHR was calculated by subtracting the average HR elicited during the final two minutes of unloaded cycling from the steady-state HR at the end of

the submaximal test. The ΔHR and ending submaximal WR were then used to generate a $\dot{V}O_2$ max estimation measured in $ml \cdot kg^{-1} \cdot min^{-1}$.

Using the test data from the female subjects, a delta HR nomogram to estimate the $\dot{V}O_2$ max of female subjects will be developed in the manner of Banister and Legge. However, no $\dot{V}O_2$ max estimations will be reported in this publication.

Statistical Analyses. Paired t-tests were used to determine if significant mean differences existed between: 1) treadmill and bike measured $\dot{V}O_2$ max values; 2) bike measured $\dot{V}O_2$ max values and the $\dot{V}O_2$ max estimations generated by the B-L nomogram; and 3) treadmill measured $\dot{V}O_2$ max values and the $\dot{V}O_2$ max estimations reported from the USAF+CF, USAF, SS+CF, and SS equations. Pearson Product Moment Correlation Coefficients (r) were generated to determine the relationship between the measured and estimated $\dot{V}O_2$ max values. Standard error of estimate (SEE) associated with the $\dot{V}O_2$ max estimations was calculated using the following formula : SEE $= S_y\sqrt{1-r^2}$, where S_y is the standard deviation of the measured treadmill or cycle $\dot{V}O_2$ max. Total error (E) was calculated using the following formula: $E = \sqrt{\Sigma(y^1-y)^2/n}$, where y^1 = measured criterion value and y = estimated value.

RESULTS AND DISCUSSION

$\dot{V}O_2$ max cannot always be measured, so accurate methods of estimating $\dot{V}O_2$ max are necessary. Submaximal tests to estimate $\dot{V}O_2$ max have traditionally been designed to estimate either a treadmill or a cycle max value. Treadmill $\dot{V}O_2$ max values in this study were significantly higher than cycle max values for both males (11.9%) and females (7.3%). These results are comparable to other studies evaluating both treadmill and cycle $\dot{V}O_2$ max values (7,11).

Baseline characteristics and performance variables for all subjects are presented in Table 1.

Table 1: Physiological And Performance Variables

Variables	Males	Females
Age, yr.	22.9±3.1	21.4±3.0
Wt, kg	78.6±13.7	59.0±8.8
Ht, in	69.1± 2.6	64.2±2.2
$\dot{V}O_2$ max TM, ml/min	4182.6±647.4	2672.5±516.8
$\dot{V}O_2$ max TM, $ml \cdot kg^{-1} \cdot min^{-1}$	53.7±6.5	45.4±5.9
HRmax TM, beats/min	198.8±8.1	194.8±8.4
$\dot{V}O_2$ max (Bike),ml/min	3667.6±505.7	2468.6±518.4
$\dot{V}O_2$ max (Bike),$ml \cdot kg^{-1} \cdot min^{-1}$	47.3±6.5	41.9±5.8
HRmax (Bike), beats/min	191.0±8.8	192.6±8.1
WRmax, W	247.6±39.5	171.4±39.8

Values are means ± SD; WR = workrate; W = watts

USAF $\dot{V}O_2$ max Estimations

Table 2 presents mean $\dot{V}O_2$ max estimations for both male and female subjects from the USAF and SS tests. The USAF tests generated significantly higher $\dot{V}O_2$ max estimations ($p<0.01$) than both SS $\dot{V}O_2$ max estimations. The USAF and SS estimations were significantly higher than the USAF+CF and SS+CF for the females subjects ($p<0.01$), whereas, the USAF+CF and SS+CF were significantly higher than the USAF and SS estimations for the males ($p<0.01$).

Table 2: $\dot{V}O_2$ max Estimations from the USAF and SS Tests

	Males	Females
Estimate	$\dot{V}O_2$ max, $ml{\cdot}kg^{-1}{\cdot}min^{-1}$	$\dot{V}O_2$ max, $ml{\cdot}kg^{-1}{\cdot}min^{-1}$
USAF+CF	48.5±8.1	41.4±7.8
USAF	47.1±9.9	45.6±10.0
SS+CF	45.7±8.2	37.9±8.1
SS	43.6±10.0	41.2±10.3

Values are means ± SD

The USAF prediction equation is designed to produce a $\dot{V}O_2$ max value which most closely estimates the $\dot{V}O_2$ max measured during a treadmill test. Therefore, statistical comparisons between treadmill measured $\dot{V}O_2$ max values and $\dot{V}O_2$ max estimations from the USAF and SS tests were completed for all subjects combined, male subjects (Table 3), and female subjects (Table 5).

The $\dot{V}O_2$ max estimates for all subjects combined using the USAF+CF and USAF prediction equation were significantly different than measured treadmill $\dot{V}O_2$ max values. The USAF+CF underestimated the $\dot{V}O_2$ max by a mean of 4.6 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$. The SEE and E associated with estimating $\dot{V}O_2$ max was 4.5 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ and 6.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$, respectively. The USAF equation underestimated the $\dot{V}O_2$ max by a mean of 3.2 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$. The SEE and E associated with estimating $\dot{V}O_2$ max was 5.6 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ and 8.0 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$, respectively.

There were significant differences between the measured and estimated $\dot{V}O_2$ max values for male subjects using the USAF and SS test data. As can be seen in Table 3, all four $\dot{V}O_2$ max estimations significantly underestimated the measured $\dot{V}O_2$ max value.

Table 3: Comparison of USAF and SS $\dot{V}O_2$ max Estimations to Treadmill Measured $\dot{V}O_2$ max Values for Male Subjects.

			Males				
Measure	Mean Difference*	t-Value	p-Value	r	SEE	%SEE	E
TM - USAF	6.6	4.08	0.0006	.66	4.9	9.1%	9.8
TM - USAF+CF	5.2	3.87	0.001	.67	4.8	8.9%	8.0
TM - SS	10.1	6.485	<.0001	.71	4.6	8.6%	12.2
TM - SS+CF	8.0	6.281	<.0001	.61	5.1	9.5%	9.8

* Values are measured $\dot{V}O_2$ max - estimated value

The USAF+CF assessment underestimated the treadmill measured $\dot{V}O_2$ max of males by a mean of 5.2 $ml \cdot kg^{-1} \cdot min^{-1}$ and the USAF prediction equation without the CF underestimated males by a mean of 6.6 $ml \cdot kg^{-1} \cdot min^{-1}$. Previous validation studies of the USAF assessment (see Table 4), which did not use the CF, also reported the assessment to underestimate $\dot{V}O_2$ max values of male subjects. The mean differences between estimated and measured USAF $\dot{V}O_2$ max in this study are less than those reported by Hartung (8) and Williford (12), but are higher than those reported by Pollock (11). The SEE of the current study is similar to the value reported by Hartung, but higher than the values reported by Williford and Pollock. The Williford and Pollock studies have much larger sample sizes and wider age ranges. Therefore, the reduced SEE in this study and the study by Hartung may be due to homogenous sample populations and may have affected the reported results.

Table 4: Summary of Previous USAF Validation Studies

Study	Mean $\dot{V}O_2$ max	Sample Size	Age	r	SEE	%SEE	E
Hartung (Male)	50.5	22	37	0.95	4.25	8.4%	-----
Williford (Male)	48.0	50	39	0.74	4.8	14.4%	-----
Pollock (Male)	48.0	67	38	0.85	6.7	14.0%	7.9
Hartung (Female)	36.6	37	32	0.76	4.1	13.5%	-----
Pollock (Female)	33.2	67	39	0.84	5.5	16.6%	6.7

Of the four $\dot{V}O_2$ max estimates generated for the female subjects (see Table 5), only the USAF estimate was not significantly different than the measured treadmill $\dot{V}O_2$ max value. The USAF+CF estimation underestimated the treadmill measured $\dot{V}O_2$ max of females by a mean of 4.0 $ml \cdot kg^{-1} \cdot min^{-1}$ and the USAF estimation overestimated females by a mean of 0.3 $ml \cdot kg^{-1} \cdot min^{-1}$. Hartung et al (7) and Pollock et al (11) have reported the USAF assessment to significantly overestimate the $\dot{V}O_2$ max values of female subjects by a mean of 3.2 $ml \cdot kg^{-1} \cdot min^{-1}$ and 2.2 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively (see Table 4). The correlation coefficients of all three studies of the USAF assessment were similar, but the current study had lower SEE and E. The subjects in this study had a much higher mean $\dot{V}O_2$ max value than the other two studies. The reduced sample size, homogenous subject population, and the difference in overall fitness levels of the subjects may all be factors in explaining the difference in study results.

Table 5: Comparison of USAF and SS Test $\dot{V}O_2$ max Estimations to Treadmill Measured $\dot{V}O_2$ max Values for Female Subjects

	Females						
Measure	Mean Difference*	t-Value	p-Value	r	SEE	%SEE	E
TM - USAF	-0.3	-0.231	0.8198	.86	3.0	6.6%	5.6
TM - USAF+CF	4.0	4.494	0.0002	.86	3.0	6.6%	5.6
TM - SS	4.1	3.494	0.0023	.92	2.3	5.1%	6.7
TM - SS+CF	7.5	9.782	<.0001	.92	2.3	5.1%	8.2

* Values are measured $\dot{V}O_2$ max - estimated value

The cross-validation study by Pollock used a much larger and more diverse population than the current study. The Pollock study contained subjects from 19 to 54 years of age and $\dot{V}O_2$ max values from 16.9 to 83.2 $ml \cdot kg^{-1} \cdot min^{-1}$. The current study, as well as the studies by Hartung and Williford, used a much more homogeneous sample. The results reported in the current study, and by Hartung and Williford, may be biased due to a small age range and similar fitness levels of the subjects. However, a consistent finding of all studies is an underestimation of males and overestimation of females. Therefore, the CF improves the $\dot{V}O_2$ max estimate for the males, but reduces the accuracy of $\dot{V}O_2$ max estimates for female subjects.

Using the best equation for males (USAF+CF) and the best equation for females (USAF) reduced the combined mean difference to 2.5 $ml \cdot kg^{-1} \cdot min^{-1}$. The SEE and E associated with estimating $\dot{V}O_2$ max was 5.3 $ml \cdot kg^{-1} \cdot min^{-1}$ and 6.9 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively. One source of error inherent to the prediction equation is the Xa factor, which is designed to estimate the $\dot{V}O_2$ at the ending submax WR. By substituting the measured $\dot{V}O_2$ at the ending submax WR and using the USAF+CF equation for men and the USAF equation for women, the $\dot{V}O_2$ max estimates were no longer significantly different. The mean difference, SEE and E were 0.3 $ml \cdot kg^{-1} \cdot min^{-1}$, 5.5 $ml \cdot kg^{-1} \cdot min^{-1}$, and 6.9 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively. Using the actual $\dot{V}O_2$ values reduced the mean difference, but the SEE and E were the same. The $\dot{V}O_2$ max estimates of the subjects' with either very high or very low maximal heart rates would be expected to affect the SEE and E the most. Improving the Xa factor would not have a meaningful impact on the accuracy of estimating $\dot{V}O_2$ max in these subjects; however, the majority of subjects would benefit from an improvement in the Xa factor. Therefore, the most accurate $\dot{V}O_2$ max estimates of this study were generated by substituting the measured $\dot{V}O_2$ for the Xa factor in the USAF equation for females and the USAF+CF for males.

SS $\dot{V}O_2$ max Estimations

The $\dot{V}O_2$ max estimates using the SS test data, both with and without the CF, were significantly different from the measured treadmill $\dot{V}O_2$ max values for both males (Table 3) and females (Table 5). Compared to the USAF estimates, the SS test consistently reported higher mean differences and E.

When comparing the USAF and SS test workrates, 9 subjects had a lower WR during the SS test, 7 had the same WR during both tests, and 26 had a higher WR during the SS test. Therefore, of the 42 subjects who completed both the USAF and SS tests, 16 subjects had the same or lower WR during the SS test. The HR responses of these subjects during the SS test should be the same or lower than their response during the USAF test. However, only 3 of the 42 total subjects had a lower HR response during the SS test. Also, for the 7 subjects with the same WR during the USAF and SS test, the HR response during the SS test averaged 9 bpm higher than during the USAF assessment. These results indicate there may be a difference between the two tests which is affecting the HR response.

The USAF and SS tests use different pedaling frequencies (50 and 75, respectively). Several studies have reported pedaling frequency to affect HR response during cycle ergometry (2,6). These studies report a higher pedaling frequency is associated with a higher HR response at the same WR. Therefore, using the USAF prediction equation, which was developed using 50 rpm, to generate $\dot{V}O_2$ max estimates from the SS test data does not appear to be appropriate. A regression equation to predict $\dot{V}O_2$ max will be developed using the responses of the 42 subjects during the SS test.

Thirty-three of the 42 subjects in this study had either the same or a higher WR during the SS test compared to the USAF assessment. Pollock et al (11) reported the USAF equation to be more accurate at estimating $\dot{V}O_2$ max when the subject completed the assessment with a higher WR, so with a new prediction equation the SS test should produce more accurate $\dot{V}O_2$ max estimates than the current USAF assessment.

Banister-Legge $\dot{V}O_2$ max Estimations

The B-L nomogram is designed to produce a $\dot{V}O_2$ max estimate for a cycle measured $\dot{V}O_2$ max. Therefore, statistical comparisons between cycle measured $\dot{V}O_2$ max values and B-L nomogram $\dot{V}O_2$ max estimations were completed for male subjects (Table 6). The B-L nomogram to predict $\dot{V}O_2$ max was developed using the submaximal HR and WR responses of 15 trained and 10 untrained, 20 to 29-year old males. The nomogram was then validated using 5 trained, 4 moderately trained, and 5 untrained subjects of the same age range. The B-L nomogram is

designed to produce a $\dot{V}O_2$ max estimate for a cycle measured $\dot{V}O_2$ max. Therefore, statistical comparisons between cycle measured $\dot{V}O_2$ max values and B-L nomogram $\dot{V}O_2$ max estimations were completed for male subjects (Table 6). All three of the methods used to estimate $\dot{V}O_2$ max (refer to methods section for explanation) significantly underestimated the cycle measured $\dot{V}O_2$ max.

Table 6: Comparison of B-L $\dot{V}O_2$ max Estimates to Measured $\dot{V}O_2$ max.

Measure	Mean Difference*	t-Value	p-Value	r	SEE	%SEE	E
Bike - BL#1	7.3	6.989	<.0001	.708	4.6	9.7%	9.2
Bike - BL#2	6.0	4.638	<.0001	.586	5.3	11.2%	9.3
Bike - BL#3	2.6	2.811	0.0086	.747	4.3	9.1%	5.6

* Values are measured $\dot{V}O_2$ max - estimated $\dot{V}O_2$ max

There were 5 trained and 26 untrained subjects identified using the ΔHR-$\dot{V}O_2$ relationship (BL#1). There were 11 trained and 20 untrained subjects identified using the ΔHR-%$\dot{V}O_2$ relationship (BL#2). All subjects were designated as trained for the BL#3 estimations. All three $\dot{V}O_2$ max estimates were significantly different than the measured $\dot{V}O_2$ max values ($p<0.01$). The combined mean $\dot{V}O_2$ max values, correlation coefficient, and SEE of all subjects in the original B-L validation group were 55.3 $ml \cdot kg^{-1} \cdot min^{-1}$, 0.98, and 2.4 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively. The subjects in the current study had a mean cycle $\dot{V}O_2$ max of 47.3 $ml \cdot kg^{-1} \cdot min^{-1}$. Therefore, the B-L nomogram was developed using a much more fit population.

The BL#1 and BL#2 $\dot{V}O_2$ max estimates were significantly different than measured $\dot{V}O_2$ max. BL#1 and BL#2 exhibited reduced correlation coefficients and increased SEE when compared to the original validation of the B-L nomogram. Although the results were much different than the original validation study, the results of the BL#1 and BL#2 were similar to the values reported for the USAF and USAF+CF $\dot{V}O_2$ max estimates (Table 3). The ΔHR nomogram, like other submaximal tests, must predict to a HRmax value (ΔHRmax in this case). As reported in the Maximal HR section below, the accuracy of estimating $\dot{V}O_2$ max of subjects' with maximal heart rates which are either significantly below or above predicted HRmax range is greatly reduced. Therefore, to have a 0.98 correlation coefficient and SEE of 2.4 ml/kg, the subjects of the original validation study must have had very similar maximal heart rates. The results observed in this study appear to be similar to other methods of estimating $\dot{V}O_2$ max, so the original validation study may have been biased by the subject population.

Although the BL#1 and BL#2 estimations were significantly different than measured $\dot{V}O_2$ max, when the $\dot{V}O_2$ max estimates are separated by trained and untrained estimations, the trained

estimations are not significantly different than the measured $\dot{V}O_2$ max values. The majority of subject are designated untrained by both the BL#1 and BL#2 criteria. Therefore, the overall accuracy of the B-L estimations may be negatively biased by the subject population of the current study. The BL#3 estimations improved the accuracy of estimating, but the estimations were still significantly different. Therefore, the use of the trained equation for all subjects is not valid. There is still a need to differentiate between trained and untrained subjects, but these results indicate the B-L nomogram is less accurate when predicting subjects of lower fitness levels..

The ΔHR-$\dot{V}O_2$ and ΔHR-%$\dot{V}O_2$ max relationships turned out to be very poor methods of stratifying individuals by training status. Of the individuals with the five highest $\dot{V}O_2$ max values, only two were classified as trained. Also, the ΔHR-%$\dot{V}O_2$ max relationship identified three subjects with $\dot{V}O_2$ max values below 40 $ml \cdot kg^{-1} \cdot min^{-1}$ as trained individuals. Submaximal tests are designed to be implemented when direct measurements of oxygen consumption are not available. Therefore, the use of the ΔHR-$\dot{V}O_2$ and ΔHR-%$\dot{V}O_2$ max relationships cannot be used outside of the research setting to designate a subject's training status. Also, the methods to identify a subject as trained or untrained were generated from the responses of only a few individuals. It does not appear these equations are applicable to even a moderate sized group and are certainly not applicable to a population the size of the USAF.

Due to the inherent inter-individual variability of the human population, the concept of using ΔHR rather than an absolute HR response is intriguing. The B-L nomogram was developed using only 25 subjects. The results of this study indicate the B-L nomogram consistently underestimated lower fit subjects, but provided accurate estimations of higher fit subjects. Based upon the current results, the B-L protocol and nomogram would not be recommended as a method to estimate $\dot{V}O_2$ max in a heterogeneous population like the USAF. However, use of the ΔHR concept to estimate $\dot{V}O_2$ max may be practical when used with the other submaximal protocols.

Invalid Assessments

The invalid rate for the USAF assessment using the modified protocol was 2.3% (1 of 42 tests). This invalid assessment was due to a HR response which was too variable. If the previous HR parameters were applied to the current test results, the invalid rate would rise from 2.3% to 26.2% (11 of 42 tests). All 11 of these invalid tests would be due to a HR response which was too variable. This invalid rate is higher than the 16.4% reported by DeWolfe et al (4) in a study of 9,000 assessments from 5 Air Force Bases, but this rate is similar to the 28% invalid rate (57 of 207 tests) reported by Pollock (11).

The largest number of invalid assessments in the current USAF protocol are due to a HR response above 85% predicted HRmax. The number of invalid tests due to a HR response above

85% HRmax predicted was 39.7% and 79% in the study by DeWolfe (4) and Pollock (11), respectively. The change in the WR progression from 1 to 2 minute stages was designed to reduce the number of invalid assessments due to a HR response above 85% HRmax. No invalid tests occurred due to a HR response above 85% HRmax in this study. Therefore, the combination of the altered WR progression and the expanded parameters for a valid steady state HR successfully reduced the invalid rate of the USAF assessment in this study.

The invalid rate for the SS test was also 2.3% (1 of 42 tests). The one invalid assessment was due to the steady state HR exceeding 85% HRmax. If the previous parameters are applied to the SS test results the invalid rate increases to 9.5% (4 of 42 tests). The other three invalid assessments would be due to a HR response which is too variable. The majority of subjects had a higher WR during the SS test than during the USAF assessment. However, only one individual exceeded 85% predicted HRmax in the SS test. Therefore, the SS test produced higher mean workrates, which should be associated with higher $\dot{V}O_2$ max estimates, without significantly increasing the number of subjects who exceed 85% predicted HRmax.

Maximal Heart Rates

The mean measured maximal HR during the treadmill test, which was the highest HR measured for the majority of the subjects, was 197.5, 196.1 and 198.9 for all subjects, females, and males, respectively. The mean predicted maximal HR was 197.8, 198.6 and 197.1 for all subjects, females and males, respectively. There were no significant differences between the measured and predicted HRmax for males or females. Also, there were no significant differences in measured HRmax between males and females. However, measured maximal heart rates ranged from 178 to 218, whereas, predicted maximal heart rates ranged from only 190 to 202.

Although there were no significant differences between measured and predicted HRmax values, the range of actual maximal heart rates far exceeded the predicted range. The USAF assessment is designed to estimate the oxygen consumption of the subject at the ending submaximal WR (Xa factor), then by estimating to a maximal HR adjusted for age, a $\dot{V}O_2$ max estimate is generated. For subjects whose HRmax is higher or lower than predicted, the accuracy of the $\dot{V}O_2$ max estimate is reduced. Twenty year old subjects were predicted to have a HRmax of 200 bpm. As previously reported, HRmax in this study ranged from 178 to 218. The two subjects with maximal heart rates of 178 and 218 were both 20 years of age. A female subject with a HRmax of 178 had a $\dot{V}O_2$ max which was overestimated by 20% and 9% by the USAF and USAF+CF, respectively. The $\dot{V}O_2$ max of the male subject with a HRmax of 218 was underestimated by 21 and 17% using the USAF and USAF+CF, respectively.

The ability to accurately estimate $\dot{V}O_2$ max is greatly dependent upon accurate estimates of maximal HR. The HRmax estimate of 220-age was the only equation of nine published equations

evaluated in this study which produced HRmax estimates not significantly different from the measured HRmax values. However, the range of measured maximal heart rates far exceeded the range of estimated maximal heart rates. The use of 220-age is recommended for use in any new equations developed to estimate $\dot{V}O_2$ max. However, the inability to identify subjects who will fall outside of the predicted HRmax range must be accepted as an inherent source of error when estimating $\dot{V}O_2$ max using a submaximal protocol.

References

1. **Astrand, P.O. and I. Ryhming**. A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaxiaml work. *J. Appl. Physiol.* 7:218-221, 1954.

2. **Coast, J.R. and H.G. Welch**. Linear Increase in Optimal Pedal Rate with Increased Power Output in Cycle Ergometry. *Eur J Appl Physiol.* 53:339-342, 1985.

3. **Cooper, K.H.** A means of assessing maximal oxygen intake. *JAMA.* 203:135-138, 1968.

4. **DeWolfe, G.W., P. Flatten, S. Chao, R. O'Dowd, D. Richardson, M. Snedecor, and S. Constable**, *Invalid Cycle Ergometry Assessment Outcomes At Five Air Force Bases* 1997, United States Air Force Armstrong Laboratory:

5. **Flatten, P., G. DeWolfe, S. Chao, R. O'Dowd, D. Richardson, M. Hite, and S. Constable**, *Modifications to the Air Force Cycle Ergometry Protocol: Impact on Pass, Fail, and Invalid Outcomes* 1997, United States Air Force Armstrong Laboratory:

6. **Hagan, R.D., S.E. Weis, and P.B. Raven**. Effect of Pedal Rate on Cardiorespiratory Responses During Continuous Exercise. *Med. Sci. Sports Exerc.* 24:1088-1095, 1992.

7. **Hartung, G.H., R.J. Blancq, D.A. Lally, and L.P. Krock**. Estimation of aerobic capacity from submaximal cycle ergometry in women. *Med. Sci. Sports Exerc.* 27:452-457, 1995.

8. **Hartung, G.H., L.P. Krock, G.C. Crandall, R.U. Bisson, and L.G. Myhre**. Prediction of maximal oxygen uptake from submaximal exercise testing in aerobically fit and nonfit men. *Aviat. Space Environ. Med.* 64:735-740, 1993.

9. **Legge, B.J. and E.W. Banister**. The Astrand-Ryhming nomogram revisited. *J. Appl. Physiol.* 61:1203-1209, 1986.

10. **Myhre, L.G.**, *A submaximal cycle ergometry protocol for predicting aerobic capacity from the Astrand-Ryhming nomogram: Validation against measured by indirect calorimetry.* 1970, Indiana University: Human Performance Laboratory Methods Manual (unpublished):

11. **Pollock, M.L., L. Garzarella, D. DeHoyos, W. Brechue, M. Beekley, G. Werber, and D.T. Lowenthal**, *The cross-validation of the United States Air Force submaximal cycle ergometer test to estimate aerobic capacity.* 1994, Center for Exercise Science, University of Florida:

12. **Williford, H.N., K. Sport, N. Wand, M. Scharff, and D. Blessing**. The prediction of fitness levels of united states air force officers: validation of cycle ergometry. *Milit. Med.* 159:175-178, 1994.

Summary and Recommendations

One of the biggest problems with the current USAF assessment is the production of an excessive number of invalid assessments. The modified USAF assessment and the SS test dramatically reduced the invalid rate compared to previously reported values. The results of the USAF $\dot{V}O_2$ max estimates were similar to other reported values for the USAF assessment, so it does not appear these modifications affected the accuracy of the $\dot{V}O_2$ max estimations. Therefore, the adoption of the modified USAF assessment should be accompanied by a significant reduction in the invalid assessment rate. If the SS test is adopted in the future, it would be expected to produce valid test results at a rate similar to the modified USAF assessment.

Of all the $\dot{V}O_2$ max estimates for males and females, only the USAF equation for females was not significantly different than the measured criterion value. The CF used for the USAF and SS $\dot{V}O_2$ max estimations improved the estimates of males, but reduced the estimates for females. Beginning January 1998, the USAF assessment will use the USAF prediction equation for males and the USAF+CF for females. Based upon the results of this study, this action would be expected to reduce the accuracy of estimating $\dot{V}O_2$ max in the Air Force population. The use of the USAF+CF equation for males and the USAF equation for females appears to be the most accurate method of producing $\dot{V}O_2$ max estimates available to the USAF. Further improvements could be made by developing a more accurate estimation of the $\dot{V}O_2$ at the ending submaximal WR (Xa factor) and developing a more accurate estimation of the HRmax (Xb and Xc factors).

Although the SS test reported less accurate results, this appears to be due to the use of the original USAF equation which was developed using a lower pedaling frequency. When compared to the USAF assessment, the SS test sets a more precise steady state WR, uses a more comfortable pedaling speed, and resulted in a higher ending steady state WR for a majority of the subjects without increasing the number of invalid assessments due to a HR response above 85% predicted HRmax. As indicated by previous studies (11), a higher ending WR appears to improve the $\dot{V}O_2$ max estimate. Therefore, when a new prediction equation is generated to estimate $\dot{V}O_2$ max from the SS test responses, the SS test should produce more accurate $\dot{V}O_2$ max estimates than the current USAF assessment.

The B-L nomogram was much less accurate than previously reported by the original authors. The results of the B-L nomogram were similar to the USAF test $\dot{V}O_2$ max estimations. However, the B-L protocol is not a feasible method of estimating $\dot{V}O_2$ max in the Air Force population for the following reasons: consistent underestimation of untrained subjects, large WR increments between stages, high rpm's used during the test, inability to accurately determine a subject's training status, and no method available to estimate the $\dot{V}O_2$ max of female subjects.

DIRECT MEASUREMENT OF DNAPL/WATER CONTACT AREA IN THE SUBSURFACE: ONE- AND THREE- DIMENSIONAL STUDIES

Kea U. Duckenfield
Graduate Student
Department of Physical Sciences

School of Marine Science
The College of William and Mary
The Virginia Institute of Marine Science
Gloucester Point, VA 23062

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

DIRECT MEASUREMENT OF DNAPL/WATER CONTACT AREA IN THE SUBSURFACE: ONE- AND THREE- DIMENSIONAL STUDIES

Kea U. Duckenfield
Graduate Student
Department of Physical Sciences

Abstract

A technique for the direct measurement of DNAPL/water contact area in a porous medium was investigated. This technique involves quantitating interfacial tracer sorption at the interface. In the first part of the study, the technique was verified by measuring the water/glass bead contact area in a packed column. Two different tracers, Triton-X100 and SDBS, were used. The resultant contact areas were similar to each other, to a geometrically derived estimate, and to an independent BET measurement. In the second part of the study, this technique was developed further through experimentation in a three-dimensional flow system which simulates a homogeneous phreatic aquifer. DNAPL/water contact area in the flow system was measured using SDBS as the interfacial tracer. Preliminary results indicate that the entire mass of the tracer was eluted. A complete analysis of the data is in progress and publication of a journal article is planned for the near future.

DIRECT MEASUREMENT OF DNAPL/WATER CONTACT AREA IN THE SUBSURFACE: ONE- AND THREE- DIMENSIONAL STUDIES

Kea U. Duckenfield

Introduction

Dense nonaqueous phase liquids (DNAPLs) have diverse industrial and household applications; as a result, DNAPLs have been extensively introduced into the subsurface environment. DNAPLs are generally long-term, low-level contaminants; however, their concentrations in groundwater, although low, are often orders of magnitude higher than drinking water standards (Geller and Hunt, 1993).

To enable effective DNAPL remediation, researchers must accurately characterize and quantify their fate and transport processes. One such process is mass transfer of the contaminant from the DNAPL phase into groundwater. Mass transfer occurs across the DNAPL/water interface and is therefore a function of the area of DNAPL/water contact (Miller et al., 1990). Hence, in order to develop reliable mass transfer models it is necessary to quantify DNAPL/water contact area.

Several factors hinder the direct measurement or indirect estimation of contact area. These include porous media heterogeneity and the changing size and shape of DNAPL bodies over time (Miller et al., 1990). Numerous investigators have conducted studies to account for contact area indirectly. The purpose of this study was to evaluate and develop a technique of direct measurement of DNAPL/water contact area using tracer methods.

In the first part of the study, tracer experiments were devised to verify the direct measurement technique by using it to calculate solid/fluid interfacial area in a porous medium. Two column experiments were performed, each with a different tracer, to measure the solid/water contact area. Results were compared with independently determined solid phase surface area.

In the second part of the study, a tracer experiment was performed in a three-dimensional flow system in which DNAPL residual zones were present. Effective DNAPL/water interfacial area will be estimated using the results of this experiment. These areas will be compared with known residual zone locations, following excavation of the flow system.

Theory

Movement of DNAPLs Through the Subsurface. DNAPLs are more dense than, and immiscible with, groundwater. Thus, a spill tends to move down through the subsurface layers as a coherent body until a less permeable layer is reached. At this point, the DNAPL body spreads laterally until it builds up sufficient capillary pressure to penetrate the next layer. The main DNAPL body continues to move down through the subsurface until it reaches an impermeable layer. If there is enough DNAPL left in a coherent body at this point, the DNAPL will form a pool or mound on top of the impermeable layer (Schwille, 1988).

As the DNAPL body moves, residual DNAPL is left behind as "blobs" snapped off from the main body and trapped in the pore spaces (Hunt et al., 1988). The distribution, size, and shape of these blobs is determined by the physical characteristics of the porous medium (e.g. grain size distribution and pore body-to-throat size ratio) and the fluid properties of the DNAPL, such as viscosity and interfacial tension (Powers et al., 1992). These blobs form immobilized residual zones in the subsurface, which slowly dissolve into the groundwater in saturated layers and evaporate into the pore spaces in unsaturated layers (Geller and Hunt, 1993).

Residual zones are difficult to locate and remediate due to (1) the physical properties of DNAPLs which lead to flow behavior distinct from that of the groundwater (Nelson and Brusseau, 1996); (2) the small amount of DNAPL required to produce a large solute plume; and (3) the decreased efficiency over time of conventional pump-and-treat methods (Mackay

and Cherry, 1989). Even investigating the location and extent of contamination can be risky, since puncturing the impermeable layer on which DNAPL is pooled may lead to contamination of the layers below (Hayworth, 1997).

Mass Transfer Models. There have been numerous approaches to describing mass transfer between DNAPLs and water. Most of the mass transfer models rely on a linear driving force equation based on empirical arguments similar to those of Fick's first law of diffusion:

$$J = -D\frac{\partial c}{\partial z},$$

where J is the diffusive flux, D is the diffusion coefficient, and dc/dz is the concentration gradient. Similarly, for mass transfer across an interface,

$$N = k_f(c_i - c),$$

where N is the mass transfer flux, k_f is the mass transfer coefficient, c_i is the concentration at the interface, and c is the concentration in the bulk solution (Cussler, 1991). The total mass transfer is the product of the mass transfer flux and the specific interfacial area (Powers et al., 1994).

In many models it is assumed that a state of "local equilibrium" exists between the blobs and the water surrounding them (see e.g. Miller et al., 1990; Hunt et al., 1988). However, this assumption does not explain field observations of aqueous contaminant concentrations significantly below equilibrium. Furthermore, laboratory experiments suggest that this assumption holds only in limited conditions, and that mass transfer is actually rate-limited. Powers et al. (1992) hypothesize that the rate limitation may be a result of shrinking blobs, increased groundwater velocity during pumping, decreased saturation of DNAPL, and/or decreased mass fraction of the more soluble components of mixed DNAPLs.

Several rate-limited and transient mass-transfer models have been proposed to account for these factors (see e.g. Powers et al., 1994; Geller and Hunt, 1993). Borden and Kao (1992) developed a two-stage model: in the first stage equilibrium conditions hold, and in the second, mass transfer is rate-limited. Other factors possibly affecting mass transfer are the partial bypassing of residual DNAPL zones by groundwater flow (e.g. Geller and Hunt, 1993) and concentration gradients that develop within multiple-component DNAPL blobs (e.g. Powers et al., 1994).

Contact Area. In the models described above, mass transfer is given as the product of mass transfer flux and specific interfacial area (contact area between DNAPL and groundwater). However, quantifying interfacial area is a major difficulty in applying these models (e.g., Geller and Hunt, 1993; Powers et al., 1994, 1992, 1991; Miller et al., 1990; Hunt et al., 1990). There have been several different approaches to accounting for this parameter. One is to assume an ideal geometry, such as perfect spheres of DNAPL. However, assuming a liquid phase volume present as spherical drops of some uniform size underestimates the surface area available for dissolution (Hunt et al., 1988). Also, effective interfacial area increases as sphere size decreases; this assumption results in overestimation of mass tranfer; conversely, use of large spherical blobs will result in underestimating the surface area and mass transfer rate (Powers et al., 1991). A second approach is to incorporate interfacial area into a "lumped" mass transfer coefficient; that is, as the product of interfacial area and the mass transfer coefficient. Although interfacial area is unknown, this lumped coefficient can be determined experimentally. However, these values are often limited to the experimental system under study (Powers et al., 1992).

A third approach is to use some surrogate parameter or parameters that may represent the effect of interfacial area. This approach has been widely explored. For instance, some investigators have considered using DNAPL saturation as an indirect indicator for interfacial

area in mass transfer models. Berglund (1997) treated the mass transfer rate coefficient (the product of the mass transfer coefficient and the specific interfacial area) as a linear function of the residual DNAPL concentration. Gvirtzman and Roberts (1991) calculated interfacial area as a function of DNAPL saturation, based on an idealized model of pore space structure. Powers et al. (1994) found that the median grain diameter and uniformity index of the medium were better surrogate parameters than the DNAPL volumetric fraction. Alternately, investigators have used parameters describing the pore structure as a measure of interfacial area (e.g. Powers et al. 1994, 1992).

Adapting the BET Measurement Technique. The limitations and disadvantages of indirect and substitutive measurements could be avoided if a direct measurement of the DNAPL/water contact area could be made. Solid/gas interfaces have long been measured using the traditional BET technique. This involves quantifying the amount of N_2 gas that adsorbs to a solid surface and calculating the amount of surface area that quantity of gas would occupy, assuming a uniform, complete monolayer of gas molecules on the solid surface. The basic relation is described by the equation

$$a = n_m a_n N, \qquad (1)$$

where a is the specific surface area (m^2/g), n_m is the moles of gas adsorbed per gram of adsorbate, a_n is the average area occupied by a molecule of gas (or packing area, in cm^2), and N is Avogadro's number (Gregg and Sing, 1982).

Saripalli (1996) investigated the possibility of adapting this approach to the fluid/fluid interface using interfacial tracers. An interfacial tracer is an aqueous solution containing a compound which accumulates at the interface between two immiscible liquids in a uniform monolayer. Anionic surfactants were chosen as interfacial tracers because of their high water solubility, low DNAPL solubility, and affinity for interfaces due to their possession of both

hydrophobic and hydrophilic moieties, and their reduced attraction for negatively charged surfaces (Saripalli, 1996).

The concentration of anionic surfactant to be employed was determined considering the surfactant property of micelle formation. Due to their amphipathic nature, at a certain concentration in water known as the critical micelle concentration (CMC), surfactant molecules congregate into spherical structures with the hydrophobic heads together in the center of the sphere and the hydrophilic tails extending into the water. The formation of micelles thus significantly alters the partitioning behavior of the surfactant; hence, it was decided to work at roughly 30% of the CMC in order to avoid micelle formation.

To adapt Equation 1 to a one-dimensional saturated porous medium, and to allow for calculation from the results of chromatographic experiments, n_m is expressed as $(R-1)PC_o$, where R is the retardation factor of the interfacial tracer, P is the pore volume of the system, and C_o is the concentration of the input source. R is calculated from the effluent breakthrough curves for the interfacial tracer and a conservative tracer introduced into the system simultaneously. For a continuous input of interfacial and conservative tracers, assuming steady flow conditions, R follows from the area bounded by the interfacial and conservative tracer breakthrough curves and the $C^* = C_o$ line (see Figure 1). R is then the ratio of the bounded areas of the tracer of interest to the conservative tracer.

Thus, the equation for the specific fluid/fluid interfacial area is as follows:

$$a_{nw} = \frac{(R-1)PC_o N a_n}{V 10^{16}}, \tag{2}$$

where a_{nw} is the specific interfacial area (cm^2/cm^3), N is Avogadro's number, a_n is the packing area ($\mathring{A}^2$), V is the bulk volume (cm^3), and 10^{16} is a correction factor to convert $\mathring{A}^2$ to cm^2.

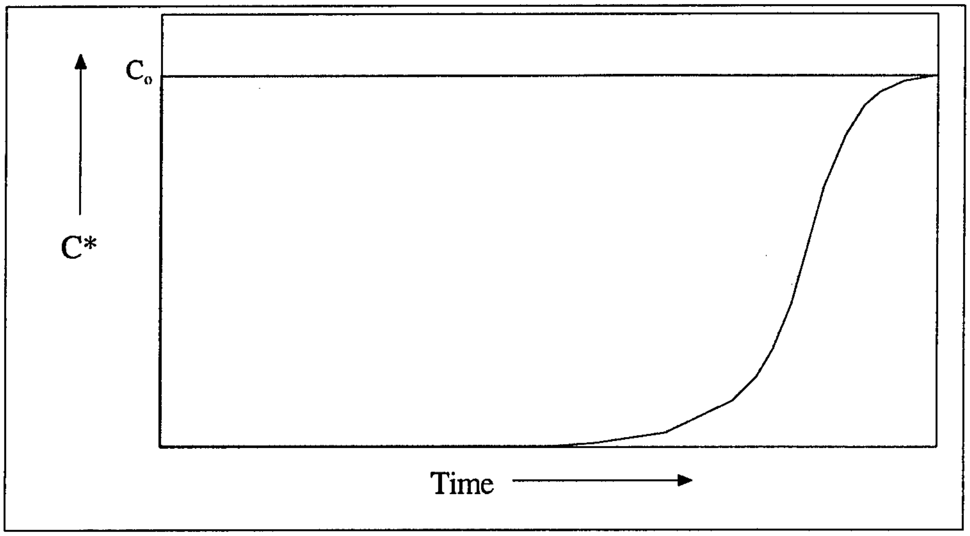


Figure 1. Schematic of an interfacial tracer breakthrough curve, with the area bounded by the line of $C^* = C_o$ and breakthrough curve indicated.

Methods and Materials

Chemicals. Two interfacial tracers were chosen: Triton-X100 and dodecylbenzenesulfonic acid, sodium salt (SDBS). Triton-X100 is an ethoxylated alkylphenol nonionic surfactant; SDBS is a linear alkylbenzenesulfonate anionic surfactant. Bromide ion in the form of aqueous sodium bromide was used as a conservative tracer in the 1-D column and 3-D flow system experiments. The DNAPL used was composed of 55% tetrachloroethylene (PCE) and 45% hexadecane, with oil red EGN used as a coloring agent.

Porous Media. Ottawa silica sand was used in the batch partitioning and box experiments. Silanized glass beads were used in the batch partitioning and column experiments.

Batch Partitioning Experiments. Batch partitioning experiments were performed using Triton-X100 to determine the partitioning behavior of the surfactant to the solid surface. Similar

experiments using SDBS are planned. Approximately 5 g of glass beads were weighed into a glass vial, and Triton-X100 dissolved in distilled deionized water (DI water) was added to the top of the vial. The final solid/solution ratio was approximately 0.25 g/mL. The concentration of Triton-X100 ranged from 5 ppm to 100 ppm. The vials were capped and placed on a rotator for approximately 48 h. The samples were then centrifuged and the aqueous phase removed using Pasteur pipettes. Final aqueous Triton-X100 concentrations were analyzed by HPLC, using a 15 cm Nucleosil C18 column, with a mobile phase consisting of 47.5% acetonitrile in DI water.

Column Experiments. A stainless steel column (2.15 cm I.D., 24.9 cm length) was packed with silanized glass beads. In preparation for each experiment, the glass beads were washed with methylene chloride, methanol, and DI water three times each. The column was rinsed throughly with DI water, dried, and packed with the glass beads, and then connected to a Waters 510 HPLC pump (see Figure 2).

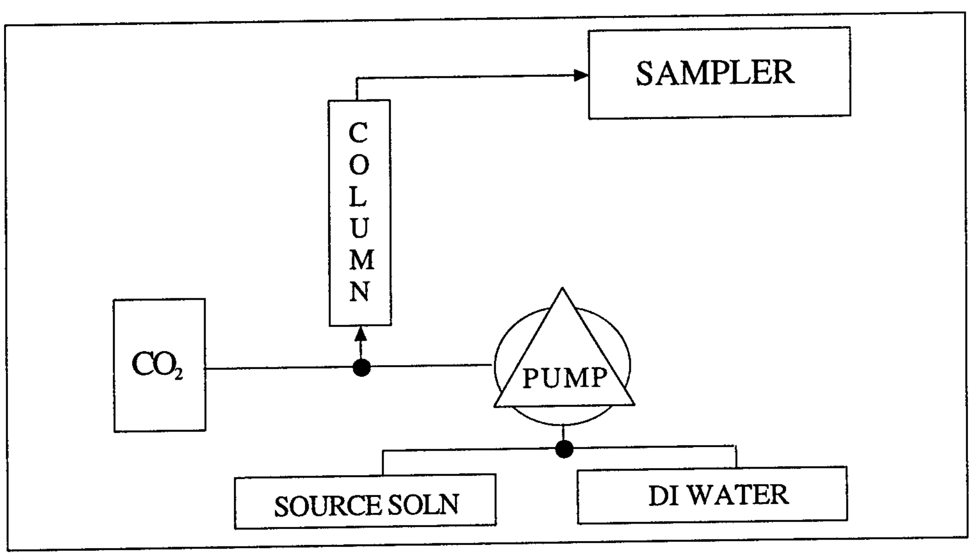


Figure 2. Schematic diagram of column apparatus.

Two experiments were conducted, one each with Triton-X100 and SDBS (Table 1). The concentration of the interfacial tracer was ~ 30% CMC. Before each experiment, the column was sparged with carbon dioxide for ~ 2 h to remove entrapped air, then saturated with CO_2-sparged DI water for ~ 2 h. At time = 0, input to the pump was switched from DI water to the tracer source. Samples were collected continuously using a Gilson 212B liquid handler. After complete breakthrough of the interfacial and Br^- tracer was achieved, the input was switched back to DI water. Samples continued to be collected until the entire mass of the interfacial tracer had been eluted.

Experiment	**Triton-X100**		**SDBS**	
Column flow rate	1 mL/min		1.2 mL/min	
Duration of pulse	35 h		18.9 h	
Experiment duration	50 h		25.6 h	
Analyte	Triton-X100	Br^-	SDBS	Br^-
Input Concentration	31.5 ppm	20.5 ppm	112.1 ppm	26.8 ppm
Sampling Interval	10 min/spl	10 min/spl	2 min/spl	2 min/spl
Instrument	HPLC	Dionex	HPLC	Dionex
Column	15 cm C18	anion exchange	15cm C18	anion exchange
Mobile Phase	47.5% CH_3CN in DI water	19mM NaOH	45% CH_3CN in 0.2M $NaClO_4$	19 mM NaOH

Table 1. Column experiment sampling and analysis conditions.

The solid/liquid contact area in the column was calculated geometrically in order to verify the proposed measurement technique. It was assumed that the beads were smooth, spherical, and uniform, and that they were in closest packing formation in the column. The empty, full (unsaturated) and full (saturated) weights of the column were measured and used to determine the porosity and pore volume of the column. BET surface area was provided by the manufacturer .

Three-Dimensional Experiment. This experiment was conducted using a three-dimensional flow system (Figure 3). The system simulates a phreatic aquifer. All wetted surfaces are constructed of non-sorbing materials, and over 500 sample ports are in place (Figure 4). The

system was packed with Ottawa sand and saturated with DI water, and a number of DNAPL residual zones were emplaced in the system (Figure 4). The flow rate at the time of the experiment was approximately 5 mL/s.

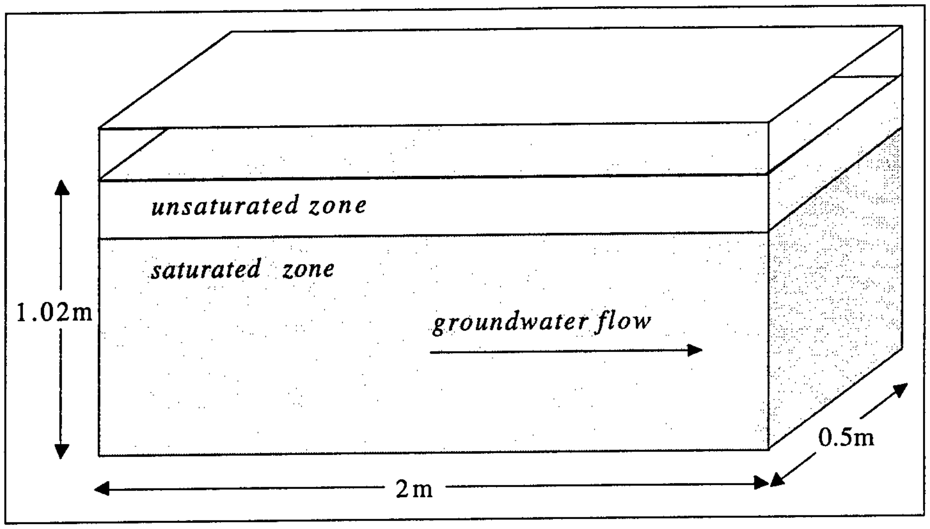


Figure 3. Schematic diagram of 3-D flow system.

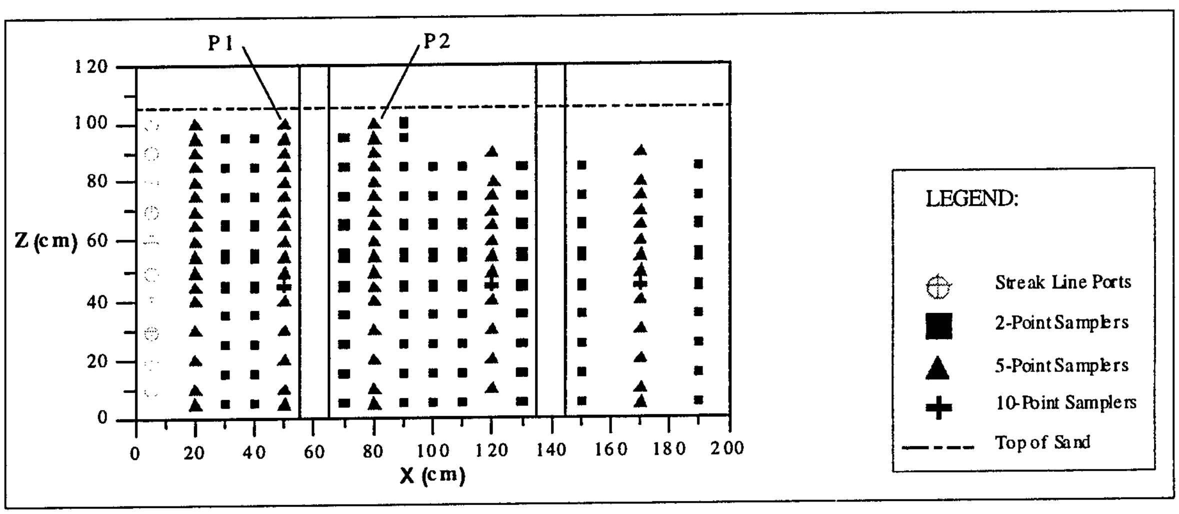


Figure 4. Schematic diagram of sampling ports for 3-D flow system. "P1" and "P2" indicate the two planes of residual DNAPL emplaced previous to this experiment. From Hayworth, 1997.

At time = 0, a one-pore volume pulse (approximately 300 L) was introduced into the 3-D flow system as a fully penetrating slug. The pulse contained aqueous bromide ion and SDBS in the same concentrations as used for the column experiment (i.e., 20 ppm and 100 ppm, respectively). Samples were taken from the upgradient and downgradient mixing chambers and eight additional points within the porous medium at 15-min to 5-h intervals as the experiment progressed. At time = 88.5 h, the experiment was terminated. Bromide and SDBS were analyzed as previously described (Table 1).

Results

Batch Partitioning. Figure 5 shows the adsorption isotherm for Triton-X100; the isotherm is nonlinear. The SDBS partitioning experiments are still in progress at the time of this report; previous work by Saripalli (1996) in similar systems suggests that its behavior is also slightly nonlinear.

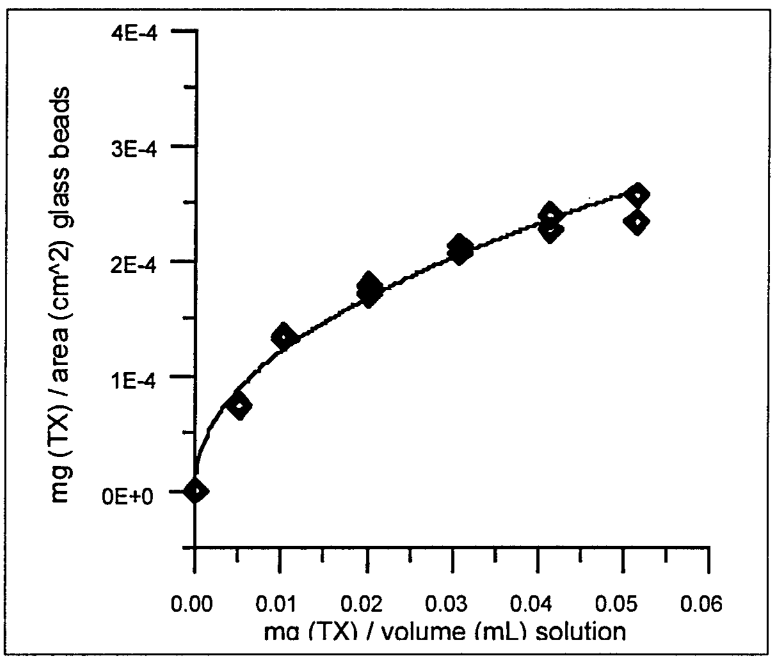


Figure 5. Adsorption isotherm for a water/Triton-X100/glass bead system. Diamond symbols represent data points; line is a power curve fit.

According to Saripalli (1996), an experiment using a pulse input of interfacial tracer exhibiting nonlinear partitioning behavior must be analyzed using a different set of equations based on the surface free energy of the contact area. However, the one-pore volume pulse may be treated as a continuous input of tracer due to the low dispersivity of the medium. In this case, the nonlinearity of the tracer does not affect the calculation of contact area using Equation 2 (Hayworth, 1997).

Column Experiment: Triton-X100. The geometric estimate of water/bead contact area in the column was 1.62 m^2. Figure 6 shows the breakthrough curves for Br^- and Triton-X100 in the column. The areas bounded by $C^* = 1$ ($C^* = C/C_o$) and the breakthrough curves were calculated by the trapezoid rule. The contact area calculated using Triton-X100 was 1.35 m^2.

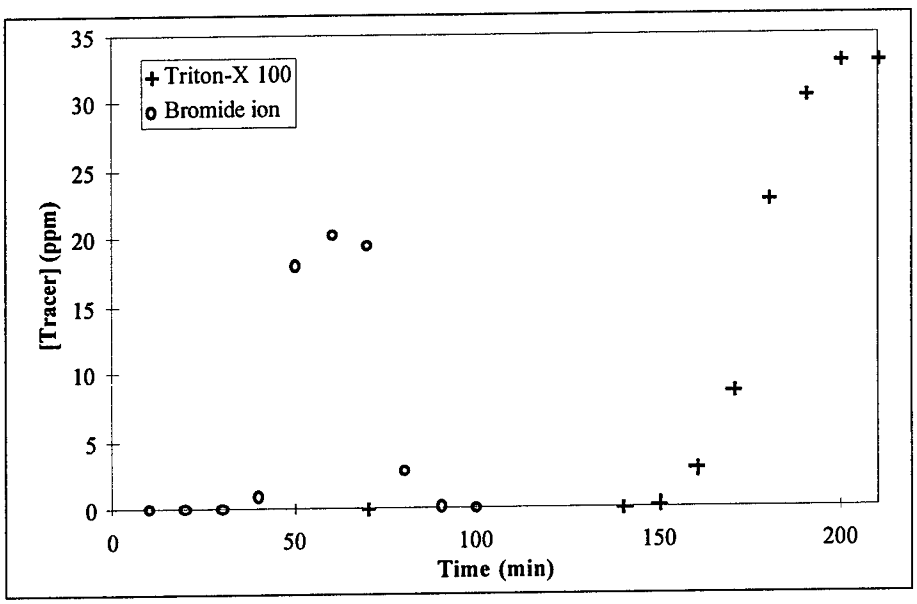


Figure 6. Results of Triton-X100 column experiment.

Figure 7 shows the breakthrough curves for Br^- and SDBS in the column. The contact area calculated was 1.41 m^2.

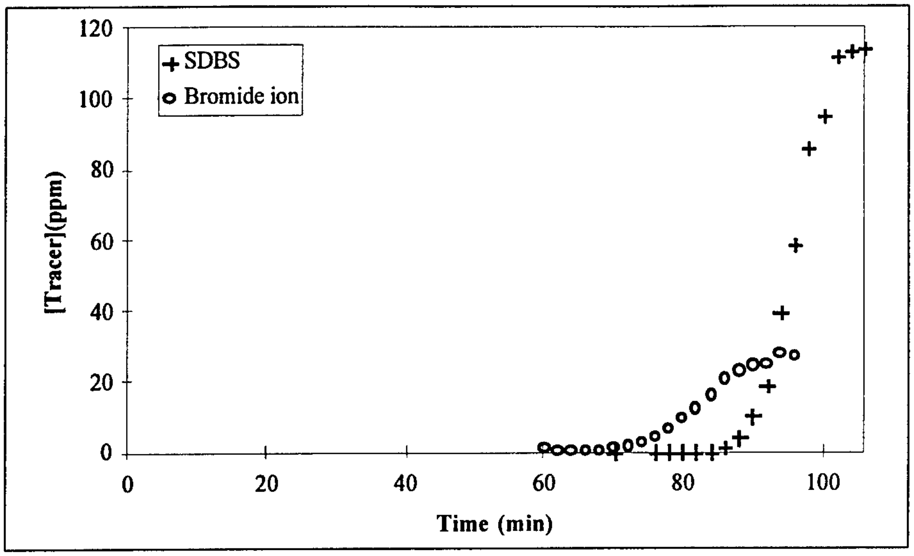


Figure 7. Results of SDBS column experiment.

3-D Experiment. Results for the 3-D experiment are shown in Figure 8. To determine whether the entire mass of SDBS was eluted, the areas under the upgradient and downgradient breakthrough curves were calculated using the trapezoid rule. If the entire amount of SDBS was eluted, the area under the curves should be the same. These areas differ by less than 1% (0.668 d vs. 0.660 d, respectively).

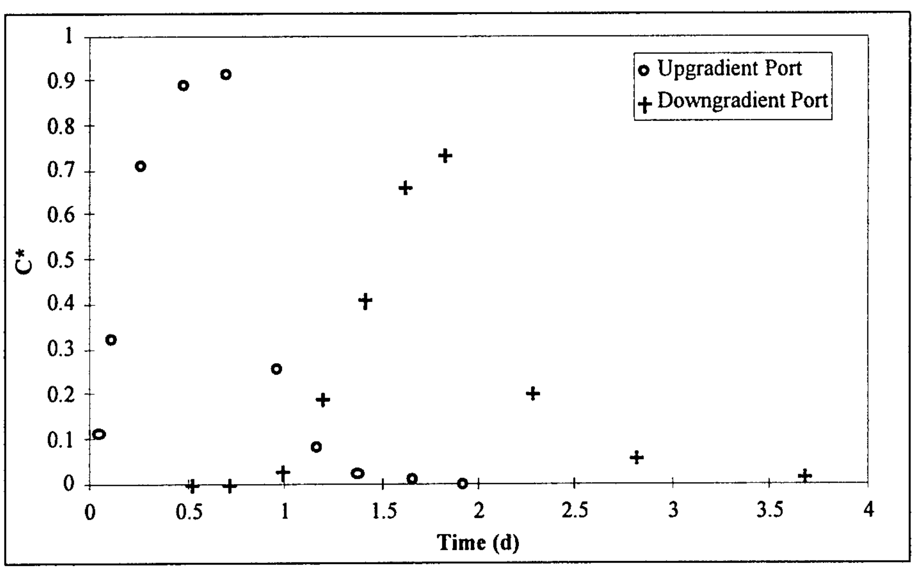


Figure 8. Comparison of the SDBS breakthrough curves for the upgradient and downgradient ports on the 3-D flow system.

Discussion

The column experiments demonstrate that the technique of measuring contact area directly with interfacial tracers is effective. In fact, the technique was effective using two interfacial tracers with different chemical properties and different sorptive behaviors. Furthermore, these results are comparable not only to the geometrically calculated area but also to the BET surface area given by the manufacturer.

The 3-D experiment appears to have been successful in that the entire mass of tracer moved through the system, allowing for a calculation of DNAPL/water contact area once the Br^- breakthrough curves have been determined and the correction for nonlinear partitioning behavior are made[1] .

Future Work.

This technique can be applied using either Triton-X100 or SDBS, and appears to work in three-dimensional systems as well as one-dimensional systems. Once calculated, the estimate of the DNAPL/water contact area in the 3-D flow system will be compared with estimates from direct measurements of the excavated source zones.

Further investigation of the transport of the interfacial tracer through the flow system will also be performed using data from the eight sampling points located within the porous medium. A field test of this technique is planned for the near future.

[1] The batch partitioning experiment using SDBS, necessary to correct for nonlinear partitioning, is slated for the near future.

REFERENCES

Adamson, A.W. (1982): *Physical Chemistry of Surfaces.* John Wiley & Sons, Inc.: U.S.A.

Berglund, S. (1997): "Aquifer remediation by pumping: a model for stochastic-advective transport with nonaqueous phase liquid dissolution." *Water Resources Research* **33(4)**, 649-661.

Borden, R.C. and C.M. Kao (1992): "Evaluation of groundwater extraction for remediation of petroleum contaminated groundwater." *Water Envir. Res.* **64(1)**, 28-36.

Cary, J.W. (1994): "Estimating the surface area of fluid phase interfaces in porous media." *J. Contaminant Hydrology* **15**, 243-248.

Cussler, E.L. (1991): *Diffusion: Mass Transfer in Fluid Systems.* Cambridge University Press, U.S.A.

Geller, J.T. and J.R. Hunt (1993): "Mass transfer from nonaqueous phase organic liquids in water-saturated porous media." *Water Resources Research* **29(4)**, 833-845.

Gvirtzman, H. and P.V. Roberts (1991): "Pore scale spatial analysis of two immiscible fluids in porous media." *Water Resources Research* **27(6)**, 1165-1176.

Hayworth, J.T. (1997): *Personal communication.*

Hunt, J.R., N. Sitar and K.S. Udell (1988): "Nonaqueous phase liquid transport and cleanup. 1. Analysis of mechanisms." *Water Resources Research* **24(8)**, 1247-1258.

Mackay & Cherry 89, "Groundwater contamination: pump-and-treat remediation." ES&T 23(6), 630-636.

Miller, C.T., M.M. Poirier-McNeill and A.S. Mayer (1990), "Dissolution of trapped nonaqueous phase liquids: mass transfer characteristics." *Water Resources Research* **26(11)**, 2783-2796.

Nelson, N.T. and M.L. Brusseau (1996): "Field study of the partitioning tracer method for detection of dense nonaqueous phase liquid in a trichloroethene-contaminated aquifer." *ES&T* **30(9)**, 2860-2863.

Powers, S.E., L.M. Abriola, and W.J. Weber, Jr. (1994): "An experimental investigation of nonaqueous phase liquid dissolution in saturated subsurface systems: transient mass transfer rates." *Water Resources Research* **30(2)**, 321-332.

Powers, S.E., L.M. Abriola, and W.J. Weber, Jr. (1992): "An experimental investigation of nonaqueous phase liquid dissolution in saturated subsurface systems: steady state mass transfer rates." *Water Resources Research* **28(10)**, 2691-2705.

Powers, S.E., C.O. Lourring, L.M. Abriola, and W.J. Weber, Jr. (1991): "Theoretical study of the significance of nonequilibrium dissolution of nonaqueous phase liquids in subsurface systems." *Water Resources Research* **27(4)**, 463-477.

Reeves, P.C. and M.A. Celia (1996): "A functional relationship between capillary pressure, saturation, and interfacial area as revealed by a pore-scale network model." *Water Resources Research* **32(8)**, 2345-2358.

Saripalli, K.P., H. Kim, P.S.C. Rao, and M.D. Annable (1997): "Measurement of specific fluid-fluid interfacial areas of immiscible fluids in porous media." *ES&T* **31(3)**, 932-938.

THE EFFECTS OF TASK STRUCTURE ON COGNITIVE ORGANIZING PRINCIPLES: IMPLICATIONS FOR COMPLEX DISPLAY DESIGN PRACTICES

Philip T. Dunwoody
Graduate Student

Department of Psychology
The University of Georgia
Athens, Georgia 30602

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

IMPLICATIONS FOR COMPLEX DISPLAY DESIGN PRACTICES

Philip T. Dunwoody
Robert P. Mahan
Associate Professor
Department of Psychology
University of Georgia

Abstract

The exponential growth of automated systems is generating demands for sophisticated display design protocols. In general, cognitive engineering practices provide display design solutions that enhance the operator's ability to manage and control complex systems. However, they often do so without adequate modeling of the cognitive system requirements for this process. In effect, while the task and display properties are well defined, the cognitive system of the user remains covert and hidden from the modeling process. This report illustrates that different cognitive organizing principles are induced by different task representations. Further, performance is dependent on the congruent mapping between task, display, and cognitive organizing principle of the operator. By externalizing the organizational principle used by an operator in a given task context, in theory one can develop representations and displays that are congruent with task and cognitive system.

THE EFFECTS OF TASK STRUCTURE ON COGNITIVE ORGANIZING PRINCIPLES: IMPLICATIONS FOR COMPLEX DISPLAY DESIGN PRACTICES

Philip T. Dunwoody and Robert P. Mahan

Introduction

The prospect of operators needing to manage greater amounts of data and information in real-time technology-based systems is continuing to drive innovative research and development strategies for creating efficient information representation protocols. However, while there is valuable work being conducted in developing theoretical frameworks for rationally guiding the design of decision support systems aimed at information rich work environments, current design strategies tend to remain haphazard and largely technology driven in nature. Innovation in sensor technology and advances in computer graphics are providing new ways to capture and represent heretofore unmeasured and unseen system data. The emergence of powerful and relatively inexpensive realtime processing capabilities (e.g., desktop SGI) is making it particularly important to understand the cognitive dynamics associated with technology based systems. For example, graphical information displays are now commonly used to encode multidimensional data structures by exploiting various perceptual organizing principles such as color, size, visual angle, depth, orientation and others (see Wickens and Carswell, 1995 for review). Future efforts will most certainly have operators working within reality systems using 3-D perspective, information flow fields and other perceptual-based features (Schiflett, 1997).

Yet, despite advances in system technology, it remains relatively unclear how features of complex systems influence the properties of the cognitive system in the user. For example, much of the display engineering research has focused on specific details of the task-display properties. These details have been well defined and rendered explicit in numerous display engineering efforts. Wickens and Andre (1990) present a compelling framework for engineering design based on a proximity compatibility relationship where spatial properties of the task are mapped to an interface in an effort to activate particular perceptual mechanisms in the operator. In the use of object-like configural displays, performance has been viewed as a category matching process where operators identify values in system parameters that are compared with learned system category states, or <u>internal model</u> defining category membership (see Wickens, 1995; Coury et al, 1989; Bennett and Flach, 1992).

However, the properties of the internal mental models that are central features to the

representation compatibility arguments have tended to remain relatively covert and hidden from the modeling process. The internalized state of the cognitive system and the fact that it is not expressed in a manner that explicates its properties limits engineering solutions because operational outcomes are a joint product of the task-display-cognitive system, where it only the task-display system that is fully understood. In effect, we still know little about the information organization principles and their relationship to information representation protocols beyond the consensus that display representations should match in some way the mental models of the users.

In an effort to generate a theoretical framework that helps externalize the cognitive system of a decision maker and thereby explicate hidden details of the task-cognition system boundary, Hammond (1980, 1990; Hammond et al, 1987) focused on the nature of information organizing principles activated by the properties of the task. The cognitive continuum theory (CCT) specifically predicts properties of the mental model of the users when engaged specific task domains. The adaptable theory asserts that task properties induce particular information organizing principles in operators that range between analytical noncompensatory models to intuitive compensatory models. A fundamental substantive contribution made by the cognitive continuum theory lies in the notion that cognitive efficiency, and thus performance, is in part, a function of the **congruence** between the properties of the task and the cognitive organizing principles employed by the decision maker . The theory essentially describes a system of two continuua- one associated with the task, and the other associated with the cognitive disposition of the individual. Oversimplifying the theory (see Hammond, 1981; Hammond et al., 1987), the assertion is that various properties of the task "**induce**" a particular mode of cognition lying somewhere between the analytical and intuitive poles on the cognitive continuum. For example, a simple, highly structured deterministic task (e.g., simple mental arithmetic) is likely to induce a mode of cognition (i.e., organizing principle) at the analytical end of the continuum. Here, the psychological/behavioral consequences of such an organizing principle is that a very proceduralized set of operations are executed, at a somewhat methodical pace with a relatively high degree of accuracy, where the subject is highly aware of the organizing principle. In contrast, a complex, ill-structured and ambiguous task is likely to induce a mode of cognition in the person that is closer to the intuitive end of the continuum. The intuitive cognitive mode would tend to be associated with a holistic organizing principle, executed quickly and with lower overall accuracy when compared to some normative standard, and where the subject would manifest less awareness of the actual

organizing principle being used in performance. In effect, the closer the congruence between the properties of the task and the model of the cognitive system given the structure of the task, the more efficient cognition is likely to be. Thus, for example, an individual utilizing analytical skills to solve a very complex ill-defined problem (and vice versa) will likely display incongruence between the organizing principle that is analytic in this case, and the properties of the task, which call for a very different organizing principle that is intuitive in nature. This incongruence will ultimately lead to inefficient cognition.

The congruence construct has emerged out of the correspondence metatheoretical orientation of Popper's three world view (1963, 1972), and has been viewed as important for successful performance by a number of researchers. For example, Hockey, (1983) refers to the matching of cognitive resources with resource demands of the task in his variable activation state theory on stress. Wickens and Andre (1990) argue the construct of proximity compatibility as the relational match between the spatial properties of the task and those of representations and displays. Tversky and Kahneman (1983) describe conditions that induce particular cognitive activities in their tests of the conjunction rule. Simon (1978, 1990) discusses how the concept of rationality must be modified in order to include intuitive elements that are compatible with an uncertain world. In these and other cases, the premise, either implicitly or explicitly stated, is that cognition must match particular features of the task.

The present study examined the congruence concept using a threat identification task where task complexity and representational format were manipulated. The hypotheses tested were guided by the nature of task induced cognitive properties defined within the framework of the cognitive continuum theory congruence principle (see table 1). In general it was hypothesized that a low complexity task structure would be best served by a numeric representational format because this task/format combination would induce a congruent analytically-based organizing principle. The simplicity of the task, its limited dimensional nature, and its objective indexing of system parameters (i.e., numerically represented) would facilitate a computational (analytical) response to task characteristics (see Hammond, 1980, 1987). Analytical cognition is demarcated by a variety of unique properties such as, sequential cue use, slow information processing, low confidence in outcomes and high confidence in process, few but large errors, high consistency, and task specific organizing principle to name a few.

In contrast, a complex task structure would favor an iconic graphic format, which would induce an intuitive-based organizing principle geared to information complexity. Here, complexity, both in number of dimensions as well as their intercorrelated statistical nature is viewed as resource intensive characteristic that has been shown an important catalyst in the induction of intuitive organizing principles (Hammond, et al., 1987; Hammond, 1990, 1996). Intuitive cognition is highlighted by a rapid form of information processing, simultaneous cue usage, high confidence in outcomes and low confidence in process, normal error distribution, inconsistent cue usage, perceptual based cue evaluation as well as others.

Method

Participants

Participants were thirty undergraduate psychology students selected from the University of Georgia Research Participant Pool or which 40 percent were female. The mean age of the participants was 23.5, SD = 2.5 ranging from 20 to 26. Participants were provided course credit for complete participation in the study. Monetary incentives were also provided to the participants with the top three performers in the study receiving $50.00, $30.00, and $20.00 for 1st, 2nd, and 3rd, respectively.

Task Overview

Research participants worked on a threat identification simulation that was adapted from the Team Interactive Decision Exercise for Teams Incorporating Distributed Expertise (TIDE^2) (see Hollenbeck et al, 1995) and modified to allow numeric and iconic protocols necessary for study, as well a provide a specific single user interface.

The software was programed to simulate a threat identification exercise where participants were presented with a number of attribute values on an object and asked to render a judgment on the status of the object based on the integration of the attribute information. Participants were taught how to weigh and integrate the attribute information in producing judgments of object state (discussed below).

The simulation is configurable and can be programed to represent a number of task domains. However, in this study the simulation was programed to represent a naval command threat assessment scenario similar to the TIDE^2 task used by Hollenbeck et al (1995). The participants were assigned the role of a aircraft carrier commander whose goal was to protect the ship from hostile airborne enemy by patrolling the airspace surrounding the carrier. In this role, the commander had access to a variety to tactical assets that provided him/her with information concerning objects appearing in the airspace.

The task was to evaluate the information on a object on the basis of attributed values such as, air speed, size, direction, angle and render a judgment as to the appropriate response to make toward that object. Judgments were rendered on a 7-point scale that varied in the intensity of the military response directed toward the object from ignore (the most benign response) to defend (destroy the target). Intermediate values of intensity or aggressiveness were also available, and included (2) review, (3) monitor, (4) warn, (5) ready, and (6) lock-on (see Hollenbeck et al, 1995 for review).

Participant Instruction. Participants were instructed on how to perform the task. First, the participants were given a very detailed verbal description of the threat assessment exercise they would be performing during the experiment. They were informed that the task was a simulation of a military command and control judgment task in which their responsibility was to patrol the air space around a carrier. Participants were instructed to rate how threatening the approaching aircraft was on a scale from one to seven based on the information provided. Each attribute was described in detail by the experimenter. For example, in the numeric display condition participants were informed that aircraft speed was represented in mph and that 100-300mph was a slow speed, 301-550mph was average speed, and 551-800mph was a fast speed. In the iconic display condition, participants were informed that aircraft speed was represented by an arrow projecting from the aircraft. A short arrow length indicated the aircraft was moving at a slow speed (100-300mph), a medium arrow length indicated the aircraft was moving at an average speed (301-550mph), and a long arrow length indicated the aircraft was moving at a fast speed (551-800mph). All attributes were described in this manner and participants were given a list of the attributes and their description that they could refer to throughout the task. After the attributes were described participants were given a description of the possible judgment values and an explanation of the feedback screen. At this point participants were informed that because this was a performance oriented study, monetary awards would be given for the top three performers. Participants were encouraged to ask questions throughout this period. After questions were answered the task was started and participants were asked to confirm that they understood the presentation of all attributes, their judgment options, and the feedback screen. This confirmation was given by participants within the first eight trials and no further questions were permitted.

Research Design. Two levels of analysis were performed on the data in test of the congruence principle. The first was an ideographic evaluation that examined each participant's judgment protocol for reliability, as well as its fit within the hypothetical framework advanced in the study. The

ideographic evaluation is part and parcel of the adaptive perspective used here where (1) judgment behavior is viewed as being mediated by a unique correspondence between the judge and the ecology in which he/she is embedded, and (2) as a result of this unique and personal correspondence, an evaluation of the reliability of the judges relationship within this ecology is tantamount to a test of the adaptive character of this relationship (see Hammond, McClelland, and Mumpower, 1980 for review).

The second level of analysis was nomothetic in nature where a 2 X3X 2 mixed analysis of variance design was used to evaluate the within factors experience and representational format manipulation which were crossed with the between factor task complexity manipulation.

Independent Variables. Three independent variables: Experience, Display, and Task Complexity were used. Experience was a three level within factors variable that provided the study with three epochs of judgments to examine. Similarly, two levels of the within subject factor representational format (Display) were used (n = 30). The first level presented the object attribute information to the participants in a numeric tabular form. Figure # shows an example of the numeric display. The information on target objects was present simultaneously on the screen in both display conditions. The second display format presented the information to the participants in a iconic graphic manner where object attribute information sources were configured as separate graphical icons, with the exception of the altitude indicator which remained numeric in nature.

Subjects were randomly assigned to each of two levels of the between subjects factor task complexity (n=15). Complexity was defined as the amount of information necessary to make judgments of target threat. In the low complexity condition participants were taught to use a decision rule that contained six attribute information values of which only four were known to the participants. In order to simulate components of a probabilistic decision task, two attributes were not visible to the participants. The invisible attributes created a barrier to perfect learning as well as perfect performance since the outcome information from each judgment was based on the entire criterion model (i.e., six information sources). The high complexity condition required participants to weigh and integrate nine information sources of which only seven were visible.

Low Complexity Criterion

$$\underline{\mathbf{Y}_{\mathbf{low}}} = (.201)X1 + (.177)X2 + (.265)X3 + (.303)X4 + \underbrace{\underline{\mathbf{(.181)X5 + (.308)X6}}}_{\text{Mystery Values}} \quad (1.0)$$

High Complexity Criterion

$\underline{\mathbf{Y_{high}}}$ =(.210)X1 + (.042)X2 +(.256)X3 + (-.027)X4 + **<u>(.191)X5 + (.398)X6</u>** +

(.124)X7 + (.411)X8 + (.110)X9 Mystery Values (2.0)

Threat attribute values were defined as follows: X1 = speed, X2 = size, X3 = Direction, X4 = Range, X5 = Radar Type, X6 = Altitude, X7 = IFF, X8 = Corridor Status, X9 = Angle of Target Vector.

<u>Performance Measures</u>. A lens model analysis of judgment task performance was conducted. The lens model can be mathematically characterized by defining the relationship among the components of the model and judgment task performance. Tucker (1964) described it as follows:

$$r_a = GR_sR_e + C[(1 - R_s^2)(1 - R_e^2)]^{.5} \qquad (3.0)$$

The correlational performance an individual achieves (i.e., achievement index) r_a, is a function of four distinct components: the linear multiple correlation between the cue values and the criterion, R_e, (environmental predictability), which indexes the uppermost predictability of the judgment task; the linear multiple correlation between the cue values and an individual's judgments of the criterion, R_s, (consistency index), which represents the ability of the subjects to control the execution of their knowledge regarding the judgment task; the extent to which the linear model of the individual judge correlates with the linear model of the criterion, G, (matching index), which measures overall task knowledge; and the extent to which the nonlinear residual variance in the model of the individual correlates with the nonlinear residual variance in the model of the criterion, designated C. Nonlinear or configural variance was negligible in this study, so the C index was not included in the analysis.

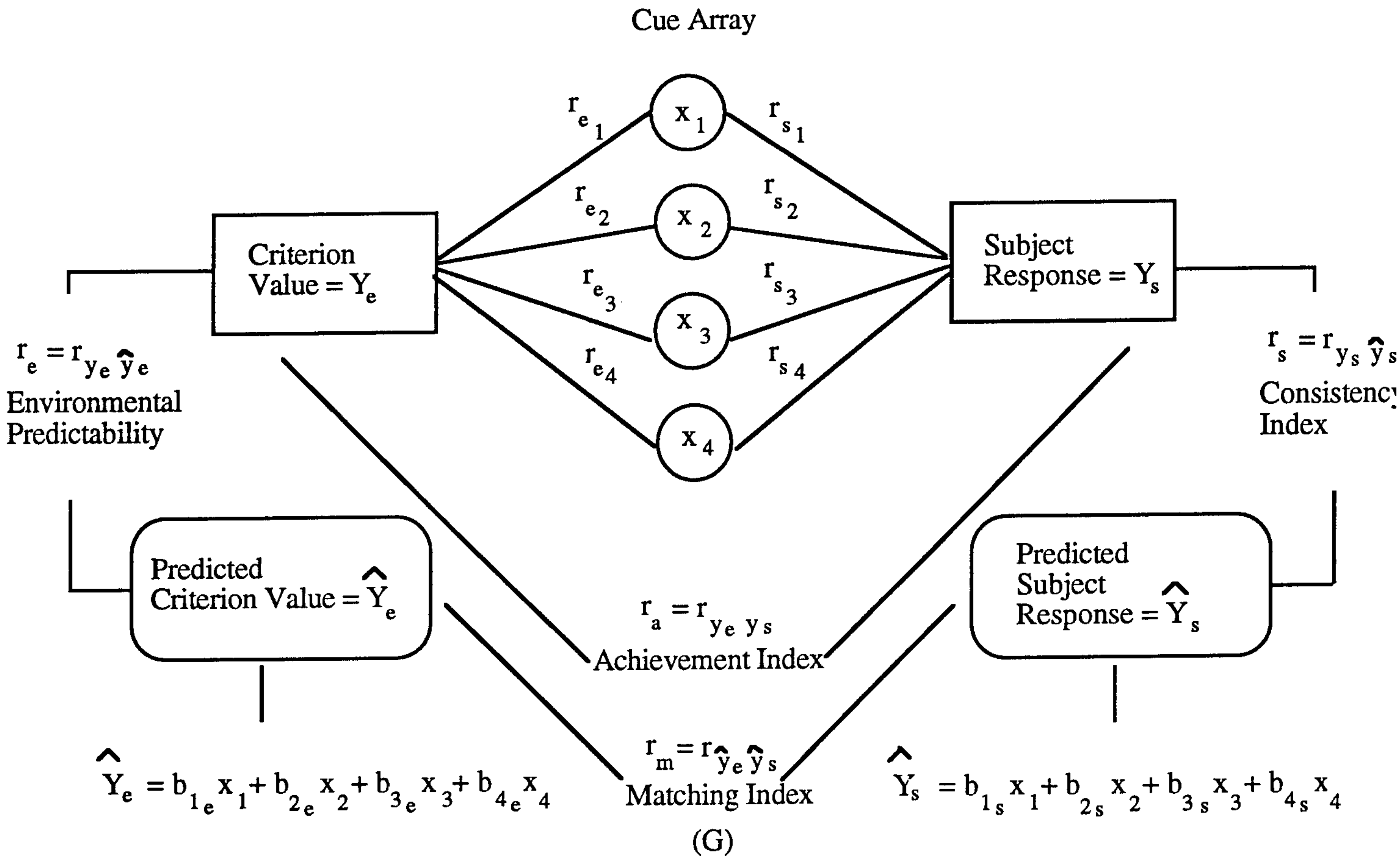


Figure 2. Formal Model of Expert Judge and Environment Interaction (From Mahan, 1992a).

Achievement (r_a) provided an index of over-all judgment accuracy. In addition, several measures that identified the mode of cognition were used. Kurtosis in the Mean Difference scores between judgment of threat identification and criterion values was used as a method to quantify error distributions for participants. A measure of implicitness toward the policy used to weight and integrate target attribute values was also computed. Here, the measure was conceived as the absolute difference between participants' subjective relative weights of attribute information and that of the least square solution of their respective policies converted to relative weights. The subjective relative eights were defined at the conclusion of the experiment where participants were asked to distribute 100 points over the attribute set in a manner that reflected the diagnostic value of each attribute in participants' judgments (Cook and Stewart, 1975). The relative weight transformation applied to the least squares policy solution was generated in a manner that allowed direct comparisons between subjective relative weights and the least squares Beta weights in the policies for each participant. The relative weighting procedure used here is a common method employed to circumvent many of the problems associated with the use of standardized regression coefficients to reflect cue importance (Anderson, 1974; Cooksey, 1996). The Beta weights in each participant's policy is divided by the Sum of the policy Beta weights to generate a set of relative weights. Implicitness then is defined as the absolute difference between the least squares policy and subjective estimates. The assertion taken was the greater the difference, the more implicit the policy.

$$rw_{\beta i} = \frac{|\beta i|}{\sum_{i=1}^{k} |\beta i|} \tag{4.0}$$

Finally, a confidence assessment was solicited from the subjects that quantified the level of belief that their subjective weights did indeed correspond to the least squares regression of judgments on target attribute arrays, or the way in which they were in reality using the information. This measure provided some insight on the knowledge over the method used to weigh the attribute information in judgments of target threat.

Procedure. Once participants were taught to perform the simulation, the study commenced. The participants performed 99 trials in each of two display formats. The formats were counterbalanced across the 30 participants. Fifteen of the participants performed the Low complexity task while fifteen performed the High complexity task. Once subjects had performed the simulation, they were debriefed. The debriefing included measuring subjective relative weights, as well as confidence rating on the subjective weights.

Results

Ideographic Descriptive Assessment

Participants were first evaluated on a case by case basis in order to determine the correspondence between outcomes and hypotheses of performance at the level of the individual. This approach is consistent with Brunswikian efforts to verify the stability (reliability) of individual indices of performance prior to nomothetic evaluations (see Hammond, et al, 1980; Cooksey, 1996).

Table 1 displays the descriptive results of the interactions between format and complexity manipulations. The focus within this table is the proportion of subjects that fall consistent with the hypotheses concerning congruence between task and display format. Five of the six outcome cells were identified as representing support for the congruence principle. The hypothesis that the numeric format should be associated with higher levels of achievement, consistency, and matching under low complexity conditions appeared supported for consistency with strong support given for matching outcomes. However, the achievement results failed to produce the correct directional outcome. Similarly, the hypothesis that the visual format would be superior at supporting higher levels of performance under the high complexity condition appeared supportive with positive outcomes for all three performance indices.

Table 1. Proportion of Participants Meeting Hypothetical Predictions for Congruence.

	LOW COMPLEXITY	HIGH COMPLEXITY
	Numeric $\geq$ Visual	Numeric $\leq$ Visual
Achievement (r_a)	*2/13*	**13/17**
		Numeric $\geq$ Visual
Consistency (R_s)	**8/13**	**10/17**
		Numeric $\leq$ Visual
Matching (G)	**12/13**	**11/17**

Nomothetic Evaluations

Factorial 2 X 2 mixed univariate ANOVA was used to evaluate average performance indices as well as indices of cognitive mode. A univariate approach was taken in order to maximize the clarity

with which the manipulations could be understood on the conceptually independent lens model indices. Table 2 displays the results of the factorial evaluation for the lens model performance indices which include power estimation values.

Table 2. Mixed Task X (Format X Subjects) Anova Table on Averaged Fisher Z Transformed Lens Model Performance Indices

Performance Index	SS	DF	MS	F	Sig	Power
Achievement (r_a)						
Within + Residual	3.67	28	.13			
Task	.19	1	.19	1.43	.242	.209
Within + Residual	1.58	28	.06			
Format	.4	1	.4	7.02	.013**	.723
Task X Format	.09	1	.09	1.67	.207	.238
Consistency (R_s)						
Within + Residual	1.57	28	.06			
Task	.03	1	.03	< 1	.49	.113
Within + Residual	.96	28	.03			
Format	.01	1	.01	< 1	.608	.057
Task X Format	.01	1	.01	< 1	.554	.066
Matching (G)						
Within + Residual	5.17	28	.18			
Task	1.5	1	1.5	8.12	.008***	.784
Within + Residual	2.52	28	.09			
Format	.65	1	.65	7.28	.012**	.738
Task X Format	1.35	1	1.35	15.03	.001***	.962

The factorial results of the Lens Model analysis demonstrated that the primary effects found in the study were associated with the achievement (ra) and matching (G) indices. Consistency (Rs) values did not produce statistically significant outcomes. Matching (G) appeared the most robust outcome with statistically significant outcomes associated with main and interaction effects. Finally, power estimates revealed that the significant outcomes were associated with adequate power values of .7 and above while the nonsignificant effects were associated with much lower power values.

Mode of Cognition

The mode of cognitive activity which in this study was used to provide evidence for a particular cognitive system state was evaluated using the same mixed factorial design that was used for the performance indices. Table 3 displays the results of the factorial analysis.

Table 3. Mixed Task X (Format X Subjects) Anova Table on Cognitive Mode Values

Cognitive Mode Index	SS	DF	MS	F	Sig	Power
Kurtosis of Error						
Within + Residual	11.69	27	.43			
Task	.06	1	.06	< 1	.716	.055
Within + Residual	12.36	27	.46			
Format	4.4	1	4.4	9.62	.004***	.847
Task X Format	1.13	1	1.13	2.46	1.28	.328
Implicitness						
Within + Residual	.77	26	.03			
Task	.83	1	.83	27.87	.000***	.999
Within + Residual	.27	26	.01			
Format	.1	1	.1	9.26	.005***	.832
Task X Format	.02	1	.02	2.07	.162	.283
Confidence						
Within + Residual	49.64	26	1.91			
Task	.12	1	.12	< 1	.804	.049
Within + Residual	23.08	26	.89			
Format	1.75	1	1.75	1.97	.172	.272
Task X Format	.68	1	.68	< 1	.389	.160

The results of the Mode analysis shows that both Kurtosis and implicitness values demonstrated statistically significant outcomes, while the confidence values failed to achieve significant outcomes. In addition, the power values appear to indicate that significant outcomes were again associated with adequate power values, while the nonsignificant outcomes were limited to, in some cases, extremely small values.

Discussion

The accurate assessment by human operators in the state of complex multidimensional systems has been directly linked to the manner in which the information about the systems is packaged and represented (Carswell and Wickens, 1987; Woods, Wise, and Hanes, 1981; Casey, 1986; Wickens and Carswell, 1995). Iconic and object oriented graphic displays appear to be a especially useful in tasks requiring information integration, particularly when the indicators of a system are intercorrelated (Goldsmith and Schvanevedlt, 1981; Bennet and Flach, 1992; Wickens and Carswell, 1995). In

contrast to correlated system dimension, systems that can be decomposed into orthogonal information dimension are thought to be better served by representations that maintain the uniqueness of the information sources through separable data displays (Goldsmith and Schvanevedlt, 1981; Coury, Boulette, and Smith, 1989; Wickens, 1986).

The aim of testing the cognitive continuum congruence construct emerges from the observation that despite advances in engineering and applied cognitive science solutions to difficult human computer interaction problems, the cognitive system of the operator often remains covert and hidden from the modeling process. This study demonstrates, in part, that (1) the idea for congruence between cognitive organizing principles and representation formats has potential in defining many of the heretofore unmeasured details of the task-cognition interaction, and (2) the congruence principle, through lens model functionalism, provides a possible means to explicate and externalize important features of cognitive system of a user.

In the present study the tasks used were differentiated by both the number of parameters and the number of significant intercorrelations among cues. The results provide some solid support for the congruence principle hypothesized to account for the effects of task and format configurations on the state of the cognitive system. The low complexity (simple) task appeared best served by a numeric display format which induced a congruent analytically anchored organizing principle in the operators. In contrast, a complex task appeared best served by an iconic display that induced a congruent organizing principle that tended to be more intuitively anchored in nature. The support generated in favor of the congruence principle is particularly interesting considering the effects have most likely been underestimated due to the constraints on methodological power.

The CCT postulates that an analytical mode of information processing is typically induced by a number of task properties including the number of cues, and the presence of the intercorrelations among cues. The low complexity criterion rule in equation 1.0 shows that decision makers were shown only four cues. Cues five and six were hidden from participants and thus called mystery cues. The mystery cues provided an element of uncertainty to the task by concealing a source of variation from participants that outcome feedback was not able to reveal. The uncertainty component was viewed as important feature in representing an operational threat assessment task where some information is likely to be absence and not available to the operators. Thus, participants could only

generate perfect augments of threat by chance alone. However, since the task required parsing only four cues and since the four cues had very low intercorrelations, an organizing principle that favored analytical decomposition of cue information by the participants would be induced by the task (Hammond, 1996).

While, the achievement (ra) values were not supported by the hypothesis that the numeric display would be superior to iconic with only 2 of 13 participants generating higher ra values under numeric condition, both consistency and matching indices did point in the hypothesized direction. The interpretation given for the achievement values, which is a measure of overall performance, suggests that the iconic display of information best supported overall performance on the low complexity task.

In contrast, consistency (Rs), which is a measure of the control each decision maker exerted over the execution of his/her decision knowledge, was highest for the majority of participants with 8 of 13 generating the largest index values using the numeric display (Table 2). The interpretation here is that the numeric display helped provide a means to consistently integrate the cue information over judgments of the threat criterion. The implication of analytical decomposition of information is that a specific rule (i.e., organizing principle) can be more easily applied to the data by the subjects in an explicit manner. Further, since the rule tends to be explicit, in part, because one is able to decompose the problem into orthogonal components, the rule is more easily applied the same way judgment after judgment (see Cooksey, 1996; Hammond, 1996; Mahan, 1994; Simon, 1983). Thus, the reliability of the decision maker's execution of judgment policy information remains high.

Similarly, the matching index values were highest for the 12 of 13 participants using the numeric display under low complexity conditions. The matching index is considered a knowledge index in the sense that in order for this index to be high, a decision maker must maintain the diagnostic rank order value of the cues that is in the model of the criterion. Thus, if a decision maker maintains this monotonic rank ordering during decision activities, he/she will generate high matching values, which is what was found for the numeric display of information. The interpretation given this finding is that the numeric display assisted participants in parsing the decision problem into four relatively orthogonal pieces that helped participants explicitly understand through feedback the weights that should be applied to each cue value.

The iconic display in the low complexity case is less effective at supporting the decision makers ability to parse individual information components because of the fact the cue values are encoded as a

perceptual feature of an object (see Wickens and Carswell, 1995 for review). Thus, the iconic display did not allow for mode of cognition that reflected the analytical decomposition of cue information.

In contrast to the low complexity task outcomes, the high complexity task presented the decision makers with a greater number of information cues that were moderately intercorrelated. Equation 2.0 shows that decision makers were exposed to a nine cue criterion rule which included a single inverted cue (negative b weight), and with the same mystery cues that were present in the low complexity criterion. Thus, the decision makers saw seven cues in the high complexity condition.

The iconic display presented cues as objects where cue values were encoded as perceptual features such as veridical size of aircraft, the perceptual distance between an aircraft and the aircraft carrier, the size of air corridor, and others. Here, the decision makers were required to perceptually measure the cue values which assisted in inducing an intuitive response to the decision problem which matched the task complexity features. Table 2 shows that the majority (13 or 17) of decision makers generated the highest achievement values using the iconic display. Thus, overall performance was best in the high complexity- iconic display combination condition for most participants.

It was also expected that the numeric display would indeed generate the highest consistency values (Rs). Here the numeric display presents decision makers with an artificially decomposed decision problem. That is, the decision problem appears as an analytic problem where an explicit analytic organizing principle can be reliably executed across judgments. However, in this case, an analytic strategy does not account for task properties such as intercorrelated cues, thus the organizing principle is incorrect. This is shown in the lower overall achievement values associated with the numeric display. While the judgments of threat are executed more reliability (consistently) with the numeric display, they are consistently <u>wrong</u>.

<u>Nomothetic Evaluations</u>

The factorial outcomes mirrored, in part, the ideographic descriptive evaluations above. A limiting constraint on the factorial analyses was the low statistical power particularly associated with interaction effects, which were integral to testing task-display hypotheses concerning cognitive state.

Average achievement on the threat identification simulation was superior for the iconic display regardless of the complexity of the task, and thus did not support the congruence principle. However, a likely reason for the failure of the congruence principle to account for achievement is the fact that the two tasks were probably not all that different. That is, while the tasks were configured to represent

positions relatively close to the poles of the task continuum, in reality it is likely they were both located midway on the continuum.

The implication is that both tasks tended to induce quasi-rational processing because both tasks were located at a midpoint on the task continuum. Since both tasks required decision makers to effectively deal with uncertainty and cue intercorrelations, the icon display was superior. The absence of differences in average consistency index values, and the fact that they remained high across formats testifies to the fact that analytical elements in the organizing principles in both tasks were present. Here, one predicts that an analytical organizing principle will tend to be executed consistently because of the explicit nature of the rule being applied to the information. The fact that both task-display configurations represented the information cues as separate and distinct components of the decision problem may have led to the decision problem being partially organized as an analytical rule.

The judgment matching index has been viewed as reflecting the subject's understanding of the properties underlying accurate task performance (Hammond & Summer, 1972; Hammond et al., 1977; Hammond, McClelland & Mumpower, 1980; Brehmer & Joyce, 1988). The matching index measures the extent to which subjects can distinguish among the cues on the basis of their diagnostic value in predicting the criterion variable. The significant interaction between task and display (Table 3, Figure 4) showed that the numeric display appeared to best match the low complexity task while the iconic display best matched the high complexity task. Here, it was evident that task understanding was a function of the congruence between task, display, and information organizing principle.

The factorial results of the implicitness (self awareness) measure mirrored the performance profile for task knowledge (Matching), and helped validate the organizing principle used by decision makers. Figure 5 shows that the level of implicitness was dependent on the task-display condition combination where the complex task generated the highest values when iconically represented and the low complexity task the smallest value when numerically displayed. These results suggest that during conditions where the depth properties of the task were complex and the surface properties were iconically displayed, the organizing principle became more implicit and less subject to conscious awareness by decision makers. In contrast, when the depth character of the task was simple (low complexity) and the surface character was defined by a numeric display, low implicitness scores were recorded. Thus, the organizing principle used under the low complexity-numeric display condition combination tended to be explicit, and consistent with a formulaic analytical response by decision makers to the threat identification task.

The error distribution found in the study, in part, supported the predictions made by the congruence principle. CCT asserts that the the application of a explicit analytical logic-based, highly retraceable organizing principle produces consistent decision behavior. Thus, identical information sources produce identical decisions and decision behavior. Hammond (1996, 1986) notes that this type of coherence-based decision behavior has a propensity for small error. That is, on average, people are able to execute the analytical formula precisely over and over (but see Klinemunz, 1985). However, when there is a error (a mistake in a parameter within the formula) a very large error results. Thus, when executing analytical organizing principles the prediction for a leptokurtic error distribution results.

In contrast to the behavior manifest from an analytical organizing principle, an intuitive principle tends to be more dynamic in nature. That is, the absence of a formula to use on the information means that the implicitly organized rule will undergo changes from decision to decision, which reflects the absence in control over the organizing principle (low consistency). However, this behavior will be manifest as oscillations around some central decision response. That is, the slight changes in the principle from decision to decision will create small deviations around a mean decision point. This produces a platokurtic error distribution that tends to be very peaked with small tails unlike the analytical counterpart (flat with large tails).

<u>Summary</u>

In many display engineering research efforts the cognitive system of the decision makers remains a mystery and covert in nature. Thus, establishing theoretical frameworks that postulate performance constructs based on matching, compatibility, and or congruence between task and display systems are difficult to achieve. Explicating and measuring properties of the cognitive system in the user will assist with theoretical efforts designed to improves and support performance in complex systems. Understanding how particular properties of the cognitive system are changed in response to changes in the depth and surface features of a task should provide information on more efficient system design techniques.

The present study in part validates the congruence principle elaborated in the CCT that the cognitive system must be congruent with the task and display systems being used. While, the study suffered from limited statistical power, the results nevertheless suggest that the congruence principle may be useful in producing coherent display design principles, and thus should undergo continued research.

Anderson, N. H. (1974). Information integration theroy: A brief survey. In D. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), Contemporary developments in mathmatical psychology, (Vol. 2, pp. 237-305). San Francisco: W. H. Freeman.

Bennet, B. B., & Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors*, *34*(5), 513-533.

Brehmer, B., & Joyce, C. R. B. (1988). Human judgment: the SJT view. Amsterdam; Oxford: North Holland.

Carswell, C. M., & Wickens, C. D. (1987). Information integration and the object display: An integration of task demands and display superiority. *Ergonomics, 30*, 511-527.

Casey, E. J., & Wickens, C. D. (1986). Visual display representation of multidimensional systems (Tech Report CPL-86-2/MDA903-83-K-0255): University of Illinois Cognotive Psychology Lab.

Cook, R. L., & Stewart, T. R. (1975). A comparison of seven methods for obtaining subjective descriptions of judgmental policy. *Organizational Behavior and Human Performance, 13*, 31-45.

Cooksey, R. W. (1996). Judgment analysis: Theory, methods, and applications. San Diego, CA: Academic Press.

Coury, B. G., Boulette , M. D., & Smith, R. A. (1989). Effect of uncertainty and diagnosticity on classification of multidimensional data with integral and separable displays of system status. *Human Factors, 31*, 551-569.

Goldsmith, T. E., & Schvaneveldt, R. W. (1984). Facilitating multicue judgments with intergral information displays. In J. Thomas & M. Schneider (Eds.), Human factors in computer systems, . Norwood, N.J.: Ablex Publishing Corp.

Hammond. (1981). Principles of organization in intuitive and analytical cognition (Rep. No. 231): Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1980). The integration of research in judgment and decision theory (Rep. No. 226): Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

MEASUREMENT OF DISPERSIVE CURVES FOR OCULAR MEDIA BY WHITE-LIGHT INTERFEROMETRY

Daniel X. Hammer
Research Electrical Engineer
Biomedical Engineering Program

University of Texas at Austin
ENS 610
Austin TX 78712-1084

Final Report for:
Graduate Student Research Program

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base DC

and

Armstrong Laboratory

September 1997

MEASUREMENT OF DISPERSIVE CURVES FOR OCULAR MEDIA BY WHITE-LIGHT INTERFEROMETRY

Daniel X. Hammer
Research Electrical Engineer
Biomedical Engineering Program
University of Texas at Austin

ABSTRACT

White-light interferometry was used to measure the wavelength dependence of refractive indices (i.e., dispersion) for various ocular components. Verification of the technique's efficacy was substantiated by accurate measurement of the dispersive properties of water and fused silica, which have both been well-characterized in the past by single-wavelength measurement of the index. The dispersion curves were measured from 400 nm to 800 nm for aqueous humor and vitreous humor extracted from bovine eyes. The principles of white-light interferometry, including image analysis, measurement precision, and limitation of the technique, are discussed. In addition, alternate techniques and past measurement of ocular dispersion are reviewed.

MEASUREMENT OF DISPERSIVE CURVES FOR OCULAR MEDIA BY WHITE-LIGHT INTERFEROMETRY

Daniel X. Hammer

1. INTRODUCTION

Recently, a technique known as spectrally resolved white-light interferometry (SRWLI) was developed to make highly accurate, real-time measurements of the dispersive properties of transparent media[1,2,3,4]. Dispersion is the physical property of a substance that describes the wavelength-dependence of the speed of light (or refractive index) as light propagates through the substance. Dispersion is present in all systems that use the refraction of light for imaging. Dispersion leads to chromatic aberration. The simplest example is a dispersive glass lens that has a wavelength-dependent focal length. The chromatic aberration becomes apparent when the lens focuses broadband white light: the focal point is spread along the optical axis according to wavelength.

The importance of dispersion measurements cannot be underestimated. They are often used to measure other physical quantities, such as pressure, density, and stress. Moreover, light propagation models must incorporate dispersion in their calculations in order to predict accurately the behavior of light. This is especially important in predicting the propagation of light through ocular media and the focusing of laser pulses on the retina. Thus, laser safety depends on an accurate measurement of dispersion for the major components of the eye. In addition, measurement of the refractive index can be used to aid in the measurement of other phenomena, for example, the temperature dependence of the refractive index, thermal lensing due to focused laser pulses, or the strength of pressure waves in water.

The refractive index of transparent liquids or solids can be measured by several different techniques. Some of the methods are based on refractometry, in which the refractive index is determined from the angle with which a beam of light exits an optical system that includes the sample. The most prominent systems in this category are the Pulfrich and the Abbe refractometers. The Pulfrich refractometer measures either the angle of emergence of rays refracted for grazing incident light with the sample placed on top of a prism, or the deflection angle of light passing through a V-shaped prism with the sample placed inside the inclined faces of the prism[5]. The Abbe refractometer measures the critical angle of total reflection[6]. These instruments operate under two limitations: the index can be measured for visible wavelengths only, and knowledge of the prism index (or the index of all other optical elements in the system) is required. Other techniques have been developed to circumvent these restrictions[7].

A second approach to measurement of refractive indices is interferometry, in which the number of fringes or the phase evolution is detected as the arms of an interferometer are moved precise distances. Several different types of interferometers have been used, including Michelson[8], Fabry-Perot[9], and Mach-Zehnder[10]. Recently, several novel optical arrangements have been developed to measure absolute refractive index with as little prior knowledge of the sample as possible, for example, without knowledge of the sample thickness, laser wavelength[11], or expansion coefficient[12]. Moreover, systems have been constructed with multiple interferometers[13] and fiber-

optics[14]. With both refractometry and interferometry, a single measurement yields a single refractive index at the laser wavelength used.

SRWLI, however, has many advantages over these techniques. SRWLI eliminates the need to use many lasers or many laser lines to make single point measurements of the refractive index over the entire visible and near-infrared spectrum. SRWLI makes a single, real-time measurement of the differential index across a band of wavelengths whose range depends upon the grating of the spectrometer. Also, the technique removes the need for complex analyses or nonlinear interpolation and extrapolation to predict the refractive index for wavelengths where measurements have not been made[15,16,17]. Lastly, measurement of the refractive index with SRWLI can be accomplished without any knowledge of the sample characteristics other than its thickness. The implications and further explanation of these advantages, as well as the technique's limitations will be expounded upon further in subsequent sections.

Although the refractive index is a very important physical parameter, there have been relatively few measurements of it for ocular media across the entire spectrum of visible and near-infrared wavelengths. The reason for this is that even a single index measurement at one wavelength can require considerable instrumentation. Several ophthalmic texts, when computing the reduced focal length or positions of the principal points of the eye, list refractive index values for ocular media[18,19]. These values are usually assumed to be at 550 nm or a similar visible wavelength. This is also true for texts and papers on comparative ophthalmology, where different species of animals are contrasted[20,21,22]. Sivak and Mandelman, who measured the index[23], and computed the chromatic aberration[24] for several species, including man, compiled the most extensive tabulation of refractive index. Other species measured include rat[25], rabbit[26], and fish[27]. In several studies of the human eye, both the transverse[28] and longitudinal[29] chromatic aberration have been measured. However, the data is sparse, and quite lacking in the NIR. The aim of this research was to remedy that deficiency.

2. EXPERIMENTAL SETUP

White-light interferometry requires an optical layout including a broadband white light source, a traditional Michelson interferometer, and an imaging spectrograph with CCD camera. The experimental setup used in this study is shown in Fig. 1. A 75-watt xenon arc lamp (Oriel Inc., Model No. 6263) was used for the broadband white light source. A collimating/condensing lens (Oriel Inc., Model No. 60008, F/0.85) internal to the lamp housing was used to maintain a relatively low light divergence. A helium-neon (HeNe) laser was aligned collinear to the white light. This allowed an easy verification of alignment as the fringes for each source were only visible within its coherence length and the HeNe laser has a long coherence length compared to the white light (~10 cm vs. 10 μm). Both laser and arc lamp were aligned through iris apertures (I1-I4) into the interferometer.

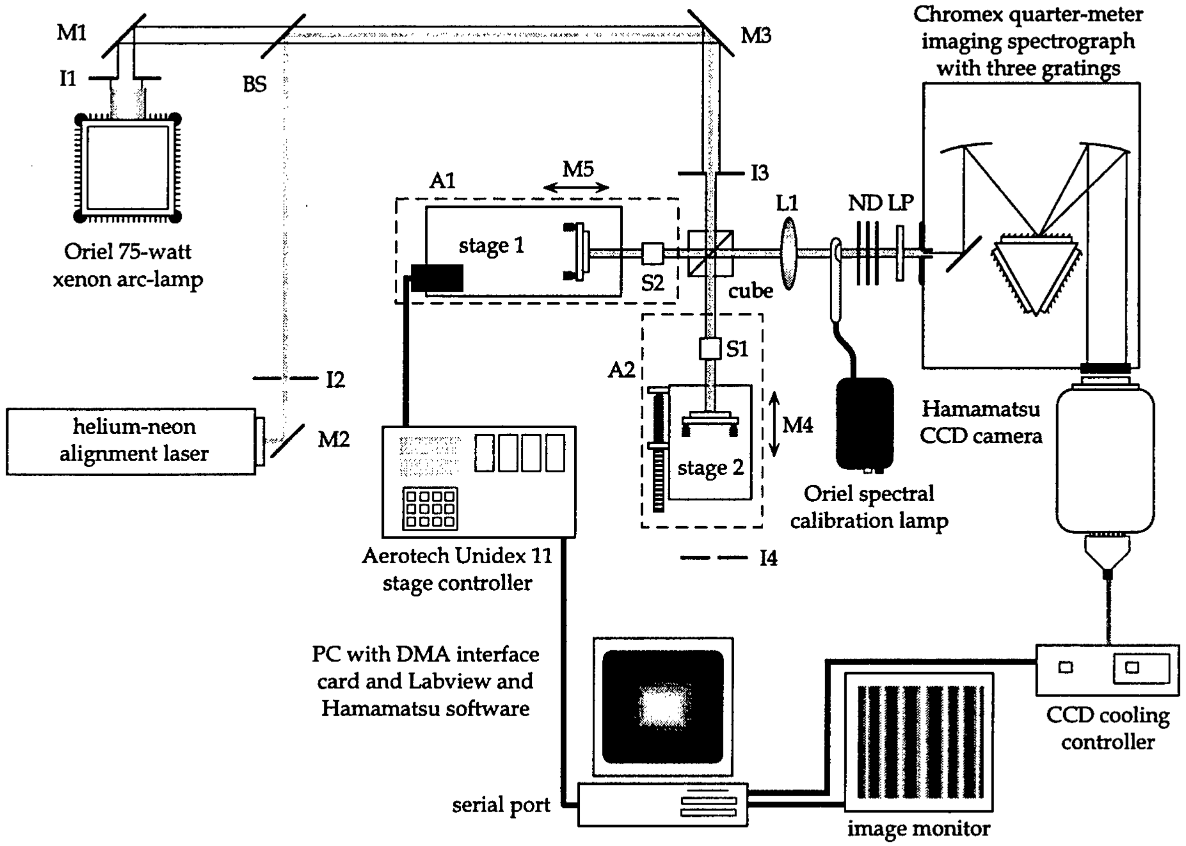


Figure 1: Optical layout for white light interferometer. A1, fixed-arm; A2, variable-arm; M1-M5, aluminum mirrors; I1-I4, iris apertures; S1-S2, sample and empty cuvette; L1, –50-mm focal length lens; ND, neutral-density filter stack; LP, 550-nm long-pass filter.

The Michelson interferometer consisted of a high-quality beam-splitting cube and two aluminum mirrors. The cube allowed the two arms of the interferometer to be balanced without a compensating plate because light travels through the same thickness of glass regardless of which arm it travels through. Aluminum mirrors were used because of their constant reflectance across a large spectral range. The fixed-arm (A1) of the interferometer consisted of a mirror mounted on a standard micrometer stage. Once gross alignment was accomplished and fringes became apparent on the image monitor, this arm was not adjusted again during the measurement. The variable-arm (A2) of the interferometer consisted of a mirror mounted on a stage with a microprocessor-based motion controller with 100-nm resolution (Aerotech Inc., Unidex 11: stage Model No. ATS1525; motor Model No. 101SMC2E-HM; stepper Model No. DM4005-400-F5; and four-axis controller Model No. U11R-2-A). The Unidex 11 stage was controlled with a Labview virtual instrument (VI) via serial line.

A cuvette containing sample was placed in one arm and an empty cuvette was placed in the other arm of the interferometer. It was not important in which arm either cuvette was placed. The cuvette thickness ranged from 2 to 10 mm, although a more accurate measurement was generally accomplished when the samples were thicker. An optical glass lens (–50-mm focal length) was placed at the output of the interferometer to expand the beam to cover the entire slit of the spectrograph. In addition, a neutral-density (ND) filter and/or a long-pass (LP) filter (85% average transmittance from 550 nm to 2000 nm) were placed at the output of the interferometer to reduce light

intensity which prevented camera saturation and blocked second-order light, respectively. The optics placed from the laser and arc lamp to the spectrograph were designed to transmit all light within the spectral range of 400 nm to 1100 nm. Optical glass lenses and ND filters limit the lower range to 400 nm, and the arc lamp and CCD camera limit the upper range to 1100 nm. (Had it been necessary to set the lower limit below 400 nm, quartz would have been used in lieu of optical glass). The spectral curve for the xenon arc lamp (after transmission through all the optics except the sample and reference cuvettes) is shown in Fig. 2.

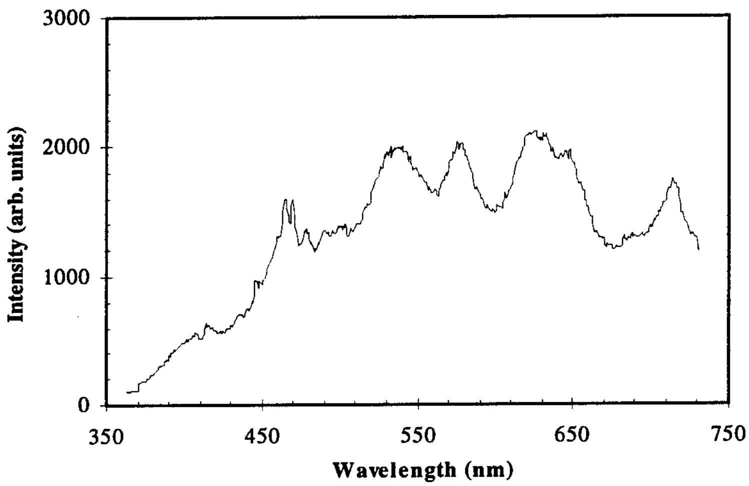


Figure 2: Spectral curve of 75-watt xenon arc lamp after transmission through the optics shown in Fig. 1.

The polychromatic light incident on the entrance slit of the quarter-meter imaging spectrograph (Chromex Inc., Model No. 250IS) was the incoherent superposition of the fringe patterns produced by the interference of each monochromatic component within the spectrum of the white light source. The spectrograph separated the intensity by wavelength, and the 12-bit 512x512 CCD camera (Hamamatsu Inc., Model No C3640-60, thermoelectrically cooled) recorded an interference image (i.e., interferogram). The image was acquired by Hamamatsu software (Temporal Analyzer Software, version 6.3) and displayed on the image monitor. The 512x512 pixel image is an interferogram with the spatial dependence (the slit length) along the vertical axis (y-axis) and the frequency (or wavelength) dependence along the horizontal axis (λ-axis). The wavelength of each image was calibrated using several rare gas spectral calibration lamps (Oriel Inc., krypton lamp Model No. 6031, neon lamp Model No. 6032, and mercury lamp Model No. 6035).

A single line (y = constant) of a typical interferogram is shown in Fig. 3. In the absence of a sample in the interferometer, the fringes will be evenly spaced in the image. The dispersion of the sample manifests itself as a modulation in the frequency of the interferogram. Thus, the dispersion can be computed by measuring the phase in the interferogram. The image analysis by which the differential refractive index was computed from the phase of the fringes of the interferogram is outlined in Section 4.

The stationary phase position is the position where the phase function is locally constant, i.e., a relative maximum or minimum. Since the phase is locally constant, the intensity is also locally constant (see Eq. (3) in Sec.

4). The stationary phase position can be observed in the images where the fringes are widest and the dispersion most apparent (see Fig. 3). It was necessary to obtain images near the stationary phase position because the nonlinear terms in the expansion of the phase function are much smaller than the linear term[30] and the nonlinear terms are those that determine dispersion. The stationary phase position should not be confused with the zero-order fringe, which is not visible because the fringe pattern is displayed in the spectral domain[30].

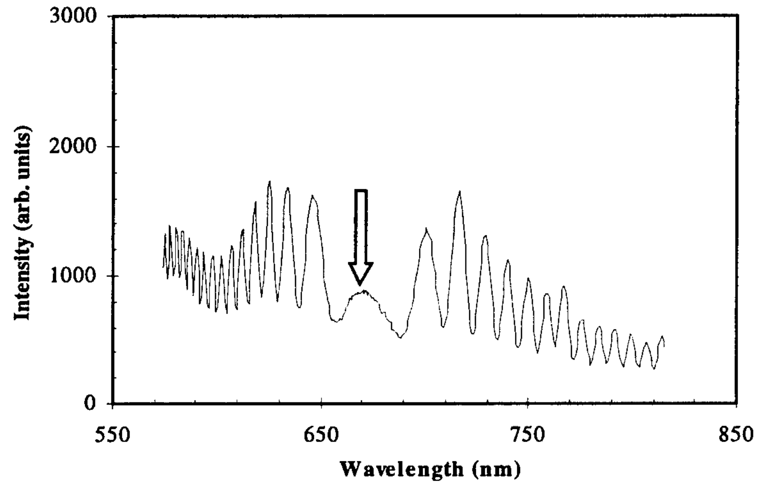


Figure 3: Single line of interferogram produced at output of imaging spectrograph for a 10-mm thick sample of high-purity water. Arrow denotes stationary phase position.

The imaging spectrograph has three gratings, two of which were used in this study: g1, 600 grooves/mm centered at 500 nm; g2, 600 grooves/mm centered at 1 μm; and g3, 100 grooves/mm centered at 450 nm. The measured window width for the 600 grooves/mm gratings is ~62 nm for a resolution of ~0.12 nm/pixel. The window width for the 100 grooves/mm grating is ~366 nm for a resolution of ~0.71 nm/pixel. The grating useful range, defined to be the region where the grating efficiency maintains at least 80% peak efficiency, is estimated to be $2\lambda_0/3$ to $3\lambda_0/2$, where λ_0 is the blaze wavelength[31]. (The blaze angle is the angle the grooves make with the back of the grating; it primarily determines grating efficiency). The specifications for the three gratings of the spectrograph are summarized in Table 1.

The gratings used in this study were g3 and g1 for visible and near-infrared (NIR) wavelength regions, respectively. However, since g3 has the largest wavelength window, it was easiest to locate the stationary phase position with this grating for both visible and NIR regions. The stationary phase position can cover quite a large spectral range, especially for the shorter wavelengths. Since the g2 grating has a higher grating efficiency in the NIR region than the g1 and g3 gratings, g2 would have been used except that the sample dispersion was too large to measure in this region. In other words, in the NIR region, the number of fringes was too small for the width of the window. Therefore, the grating with the largest window (g3) was used for the NIR region. It did not appear that the use of grating g1 outside of its useful range hindered the acquisition of correct interference patterns in the NIR region.

Table 1: Summary of spectrograph grating specifications. The useful range is estimated to be $2\lambda_0/3$ to $3\lambda_0/2$, where λ_0 is the blaze wavelength. The measured window width is shown, and the specified window width is in parentheses.

Grating No.	Blaze Wavelength (nm)	Diffraction Grating (grooves/mm)	Useful Range (nm)	Window Width (nm)	Resolution (nm/pixel)
g1	500	600	333-750	61.7 (62.5)	0.12
g2	1000	600	667-1500	61.7 (62.5)	0.12
g3	450	100	300-675	365.7 (375.0)	0.71

Another factor that can affect accurate acquisition of images is coma. Coma is an aberration due to the variation in magnification with aperture[32]. Coma manifests itself as a wing on one side of a spectral line. The effect of coma is not large up to a wavelength of 1500 nm for g1 and g2, and up to a wavelength of 6 μm for g3.

3. IMAGE ACQUISITION PROCEDURE

The procedure for image acquisition was as follows. The arc lamp and HeNe laser were grossly aligned collinear to each other through the iris apertures. Then the arc lamp was blocked and the HeNe laser was precisely aligned into the interferometer. This was accomplished by adjustment of the angle and tilt knobs on the mirror in one of the interferometer arms. It was not critical that both mirror's back-reflections were aligned, nor was it critical for the light from both arms to be aligned with the light input to the interferometer. The only critical alignment was the two arms with each other and the alignment of this superimposed light into the spectrograph. Adjustment of tilt corresponded to fringe direction and adjustment of angle corresponded to fringe width. Thus, by successive adjustment of angle and tilt, the fringes could be made vertical and wide. Adjustment of tilt ended when less than one fringe (the zero-order fringe) was visible on the entrance slit of the spectrograph.

Once the laser was aligned, the white light was allowed to pass into the interferometer and the laser light was blocked. The spectrograph grating and its position were adjusted so that the center wavelength and width corresponded to the region of interest. Then, by observation of the output of the spectrograph on the image monitor while adjusting the fixed-arm (A1), the arms of the interferometer were balanced and the position recorded with the stage alignment VI. The arms were balanced when the stationary phase position was centered on the 256th pixel. As mentioned earlier, this was more easily accomplished with the lower resolution grating (g3). After the arms were balanced, the fixed-arm was not adjusted again during the measurement. The stationary phase position in the absence of the sample was found to be nearly the same at any wavelength. Thus, it was not necessary to make multiple measurements of this position at different center wavelengths.

A cuvette holding the sample was then placed in one arm and an empty cuvette placed in the other arm of the interferometer. In the case of the fused silica, the silica was placed in one arm and the other arm was left empty. Since new optical elements were introduced into the setup, the interferometer needed realignment. This was accomplished by adjusting the angle and tilt of one mirror while viewing the HeNe fringes until the zero-order fringe was visible. Again, the HeNe was blocked and the white light allowed to pass into the interferometer. By adjustment of the distance between mirrors with the variable-arm, an image near the stationary phase position was

captured (i.e., on one side or the other of the stationary phase position). The position of the stages at the time of the measurement was recorded with the stage alignment VI and the difference between the position at the time of measurement and the position when the arms were balanced in the absence of the sample was computed. This difference was used to determine the center index for the wavelength region of interest. After the interferogram was captured, both the white light and the laser were blocked and one of the rare gas spectral calibration lamps was placed in front of the spectrograph entrance slit to acquire a second image. The gas lines in this image were later used to calibrate the interferogram wavelength.

The spectrograph position was then set to the next center wavelength, and the procedure of the previous paragraph was repeated. In this way, the entire dispersion curve could be mapped out across the wavelength region from 400 nm to 800 nm in four or five measurements.

4. IMAGE ANALYSIS

4.1. Derivation and calculation of the differential refractive index

Once an image was acquired, a second Labview VI accomplished the image analysis, based upon the derivation of the differential index by Sáinz *et al.*[30]. The optical delay, Δ, for a dispersive sample is,

$$\Delta(\sigma)=\{d[n(\sigma)-1]-L_0\}, \tag{1}$$

where n is the index, d is the sample thickness, σ is the wavenumber ($\sigma = 1/\lambda$), and L_0 is the path difference between the arms of the interferometer. The phase function, ϕ, is related to the optical delay by,

$$\phi(\sigma)=4\pi\sigma\Delta(\sigma)=4\pi\sigma\{d[n(\sigma)-1]-L_0\}. \tag{2}$$

The phase function, optical delay, and index can all be expressed as a function of the wavenumber. The intensity of light at the entrance slit of the spectrometer can be expressed as,

$$I(\sigma)=I_0(\sigma)(1+\gamma\cos\phi), \tag{3}$$

where I_0 is the intensity of the arc-lamp, and γ is the visibility of the fringes. Eq. (3) can be solved for the phase,

$$\phi(\sigma)=\cos^{-1}[(I-I_0)/I_0\gamma]. \tag{4}$$

The parameters I_0 and γ within a local region (between successive minimum and maximum) are related to the peaks by,

$$I_0 = (I_{\max} + I_{\min})/2, \tag{5}$$

$$\gamma = (I_{\max} - I_{\min})/(I_{\max} + I_{\min}), \tag{6}$$

where I_{min} and I_{max} are the intensities of the successive peaks. Eqs. (5) and (6) can be substituted into Eq. (4) to yield,

$$\phi_{\exp} = \cos^{-1}\left[\frac{2I - (I_{\max} + I_{\min})}{I_{\max} - I_{\min}}\right]. \tag{7}$$

This experimentally-determined phase will require alteration in order to solve for the differential index. The Taylor series expansion of the index is,

$$n(\sigma) = n_0 + \frac{dn}{d\sigma}(\sigma - \sigma_0) + \frac{1}{2}\frac{d^2 n}{d\sigma^2}(\sigma - \sigma_0)^2, \tag{8}$$

where n_0 is the arbitrarily chosen center index at a wavenumber, σ_0. Eq. (8) may be substituted into Eq. (2), and compared to the polynomial expansion of the phase,

$$\phi(\sigma) = A\sigma^3 + B\sigma^2 + C\sigma + D. \tag{9}$$

Solving for Δ_0 in terms of the polynomial coefficients yields,

$$\Delta_0 = d(n_0 - 1) - L_0 = \frac{C + B\sigma_0 + A\sigma_0^2}{4\pi}. \tag{10}$$

With this result, the phase can then be corrected by,

$$\phi(\sigma) = \phi_{\exp}(\sigma) - \phi_{\exp}(\sigma_0) + 4\pi\sigma_0\Delta_0 \tag{11}$$

This correction is equivalent to subtracting the constant D from the entire set of phase values. The corrected phase can be compared to the zero dispersion case, ϕ_0,

$$\phi_0(\sigma) = 4\pi\sigma_0\Delta_0(\sigma/\sigma_0) \tag{12}$$

With this result, the differential refractive index can be calculated,

$$\Delta n = (n - n_0) = \frac{\phi - \phi_0}{4\pi\sigma d} \tag{13}$$

Thus, the differential refractive index can be calculated by analysis of the phase and measurement of the sample thickness. In order to measure the absolute index, knowledge of the center index was required. Past measurements of index provided a value for the center index. However, an alternate method was also used. By recording the position of the variable-arm stage without a sample, and the position of the variable-arm stage when the measurement was taken, L_0, was computed. The center index, n_0, was then computed from Eq. (10).

4.2. Outline of analysis VI

Once the interference image was acquired, the analysis was performed as follows. A single line of a typical interferogram is shown in Fig. 4. As can be seen, the analysis region was chosen as near to the stationary phase position as possible without including it. The first step was to filter each line of the interferogram with a Butterworth low-pass filter. The output of a single line from the filter is shown in Fig. 5.

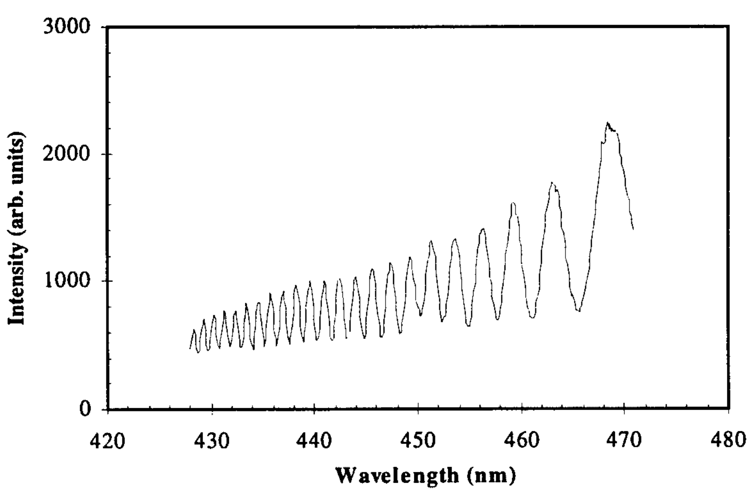


Figure 4: Single line of interferogram produced at output of imaging spectrograph for a 10-mm thick sample of bovine vitreous.

The filter removed any high-frequency noise from the curve. However, it also introduced artifact to the phase, namely a segment was added to the beginning of the curve and the entire curve was shifted toward longer

wavelengths. The added segment is unimportant because the analysis was preformed only on the portion of the curve after the first maximum. The shifted curve may introduce errors in the phase measurement. However, the shift was thought to be uniform and thus not a serious problem.

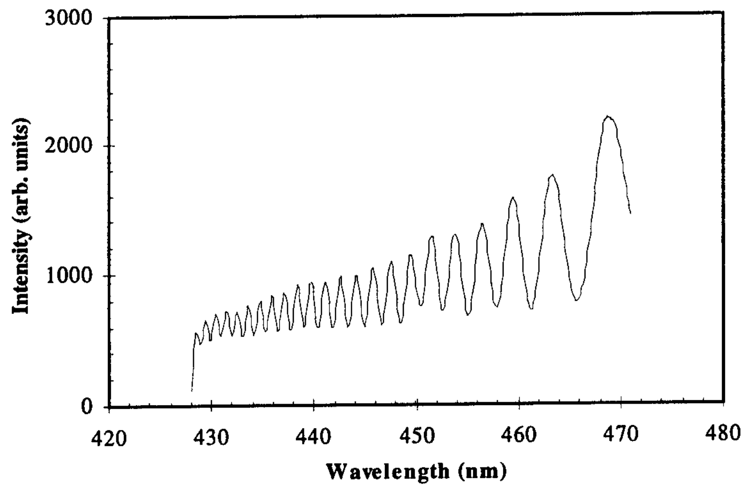


Figure 5: Interferogram shown in Fig. 3 after Butterworth low-pass filter.

Once the interferogram was filtered, the phase was computed for each segment of each line by applying Eq. (7) between every local minimum and maximum. Additive phase correction for each segment yielded a curve similar to that shown in Fig. 6. This curve was fit to a polynomial of the form of Eq. (9), and the coefficients were used to calculate Δ_0 (see Eq. (10)). The phase was then corrected according to Eq. (11), and the differential index found by Eq. (13). A typical differential index curve is shown in Fig. 7.

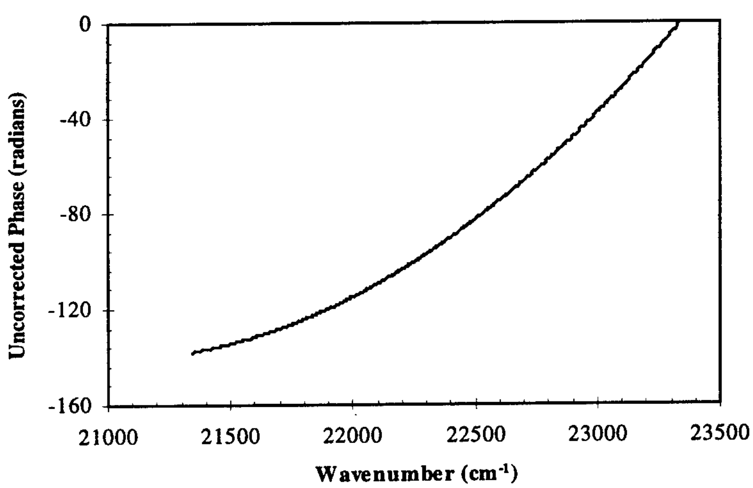


Figure 6: Uncorrected phase of filtered interferogram shown in Fig. 4 computed from analysis software.

As can be seen from Fig. 7, the differential index requires a fixed center index to give absolute index values for the sample. When the center index was found from the stage positions, the differential index at each

wavenumber could simply be added to the center index to find the absolute index. However, when previous measurements were used, the wavenumber of the index value rarely corresponded to the center wavenumber of the differential index curve. Therefore, the absolute index was set to the past index at that corresponding wavenumber, i.e., $n(\sigma)=n(\sigma_0)$. The absolute index for every other wavenumber was found by,

$$n(\sigma)=n(\sigma_0)+\Delta n(\sigma)-\Delta n(\sigma_0). \tag{14}$$

Thus the absolute index was computed from the differential index.

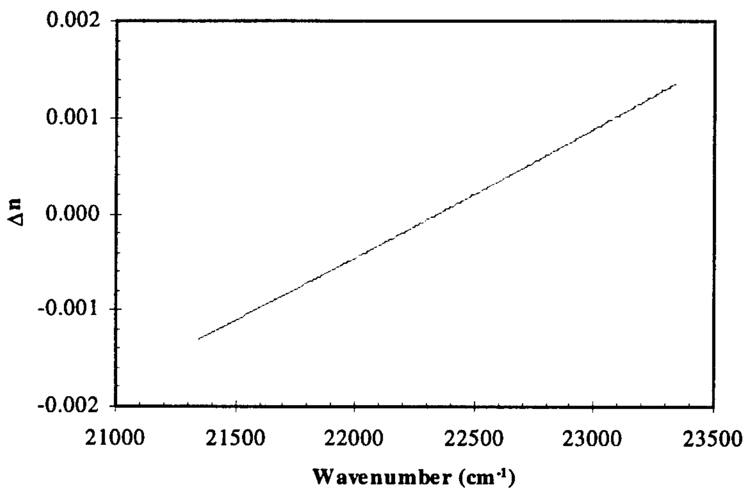


Figure 7: Differential index of filtered interferogram shown in Fig. 4 computed from analysis software.

5. RESULTS

The dispersion of 10-mm thick fused silica is shown in Fig. 8. The solid symbols are the reference data points taken from the manufacturer's literature[33]. The solid line is composed of six individual curves representing measurements of the differential index by white-light interferometry. Typically, the measured differential index was a pixel-by-pixel average of approximately 100 lines (y = constant) of the interferogram, although any number of lines greater than ~20 was sufficient for a reliable measurement. As mentioned in Sec. 4, there are two methods for computing the absolute refractive index from the differential index: by using known reference data points to compute the center index, and by using the stage positions to determine the center index for the curve. Since the error for the center index measurement was much larger than the error for the differential index measurement, the absolute refractive index was usually computed with data from past measurements. This was the case for fused silica where the difference between center index computed with past data and the center index computed with stage positions was less than or equal to 0.0063. Whereas the difference between the absolute index of past measurements (solid symbols in Fig. 8) and the absolute index computed with the differential index and the past data (solid curves in Fig. 8) was less than or equal to 0.0003.

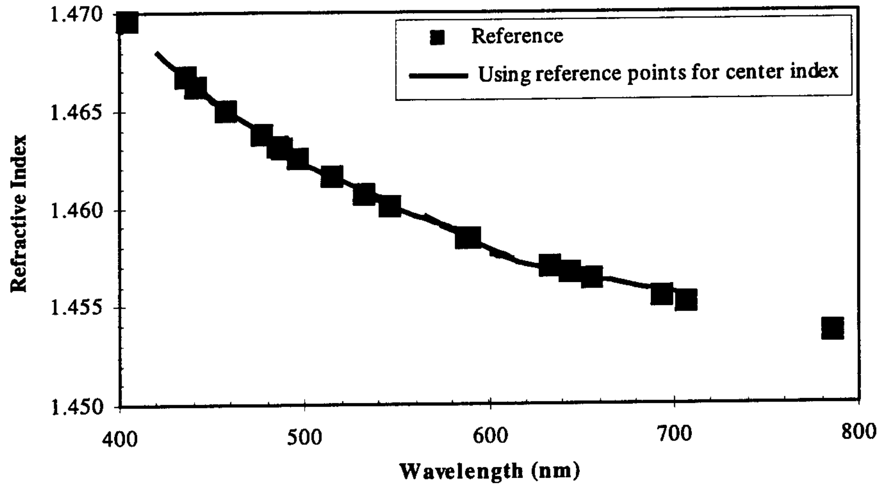


Figure 8: Refractive index of *fused silica* from 400 nm to 800 nm. Solid symbols indicate reference points from past measurements[33]. Individual curves (thick lines) are computed from analysis of the average value of 100 lines of interference pattern.

The dispersion of 10-mm thick sample of high-purity water is shown in Fig. 9. The symbols and curves represent the same measurements as those in Fig. 8, with the exception that a set of additional curves is shown (thin lines) in which the absolute index was computed from the center index measured from the stage positions (for first four sets of data only). The lines for the curves centered near 450 nm are superimposed upon one another. It is not known why the stage positions yielded reliable results for water but not for fused silica, bovine aqueous, or bovine vitreous. The difference between center index computed with past data and the center index computed with stage positions for the last two curves was larger (≤ 0.0090) than that of the first four measurements (≤ 0.0012) and so these curves were not shown. The reference data for high-purity water was taken from the fits of Stolarski *et al.*[11], who used dual-beam interferometry to make the measurements.

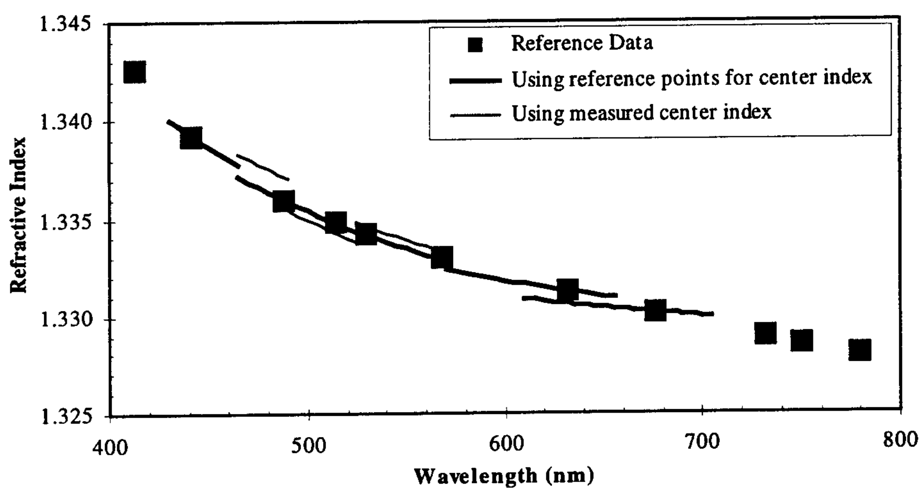


Figure 9: Refractive index of *high-purity water* from 400 nm to 800 nm. Solid symbols indicate reference points from past measurements[11]. Thick lines are measurements taken using the reference data points to set the center index. Thin lines are measurements taken using the software output to set the center index.

The dispersion of bovine aqueous humor (2-mm and 5-mm thick samples) and bovine vitreous humor (10-mm thick samples) is shown in Fig. 10 and Fig. 11, respectively. Neither figure shows the index computed from the stage positions. The difference between center index computed with past data and the center index computed with stage positions was less than or equal to 0.0058 for bovine aqueous and less than or equal to 0.0052 for bovine vitreous.

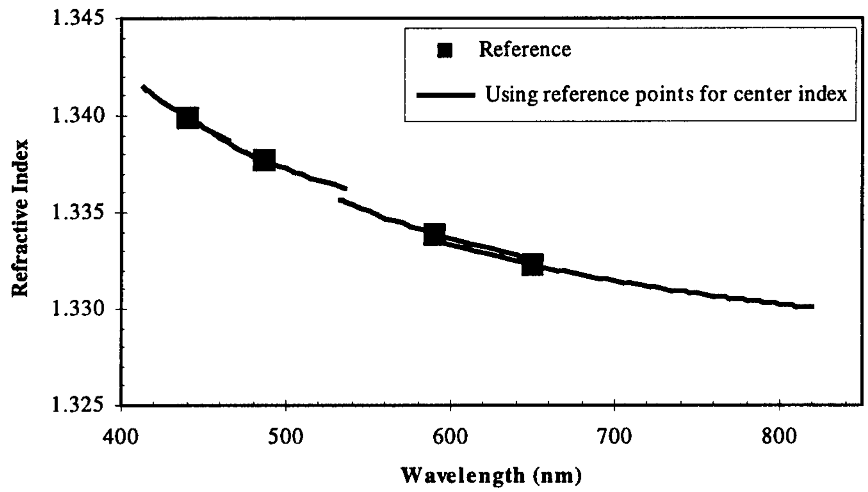


Figure 10: Refractive index of *bovine aqueous humor* from 400 nm to 800 nm. Solid symbols indicate reference points from past measurements[23]. Individual curves are differential index calculated from analysis of the average values of 100 lines of interference pattern.

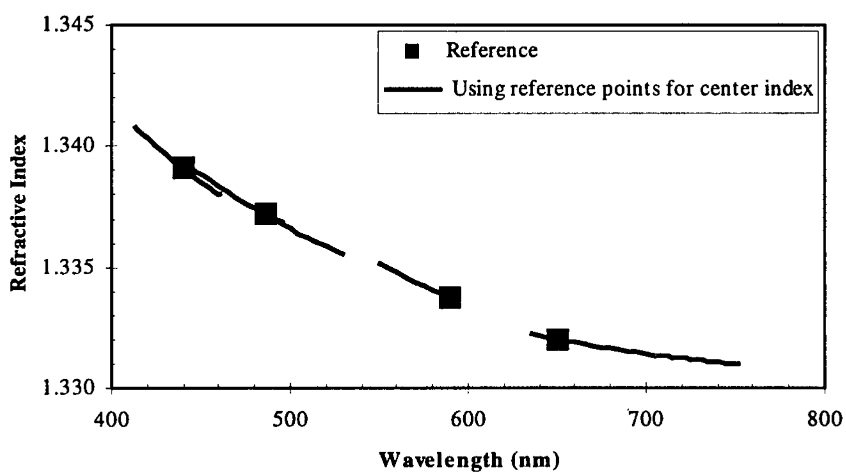


Figure 11: Refractive index of *bovine vitreous humor* from 400 nm to 800 nm. Solid symbols indicate reference points from past measurements[23]. Individual curves are differential index calculated from analysis of the average values of 100 lines of interference pattern.

6. DISCUSSION

Recently, there has been interest in an accurate measurement of the refractive index of ocular media to input into propagation models that describe the path that light takes through the eye. For example, in a model developed by Rockwell *et al.*[34], the author was compelled to use the known dispersive properties of water to predict those for ocular components out to a wavelength of 1 μm. Even though there has been some work done for visible wavelengths for some species[23], there are large gaps to be filled, especially in the NIR. White-light interferometry is a technique that has the potential to fill those gaps.

In this preliminary study, we measured the dispersion in the visible and NIR regions to establish the feasibility and accuracy of the technique. One factor that prevented measurement to 1.1 μm in this study was the xenon arc-lamp, which had several spikes from ~700 nm to 1000 nm (not shown in Fig. 2). The extraneous peaks prevented precise measurement of the phase. Further experiments are planned with a quartz tungsten halogen (QTH) lamp, which has a flat spectral output to 1.1 μm.

Of all the ocular components, it is important to find the dispersion of cornea for two reasons. First, the change in refractive index is larger between air and cornea than the change between any other two successive layers of the eye (ignoring for now the tear layer). Therefore, the refractive power of the eye is due almost exclusively to the cornea. Second, the difference in index is larger between water and cornea than between water and aqueous or vitreous. This means that water cannot be used as a substitute for cornea in the models. The differential index of bovine cornea could not be measured despite several attempts. The excised cornea often became wrinkled and always became more turbid, and this prevented measurement of sharp interferograms. Also, with this technique it was considerably more difficult to measure accurate curves when the sample was thin. Moreover, in contrast to aqueous and vitreous, which could be placed in a cuvette of known thickness, the exact thickness of the layer of cornea was not known. This precluded any attempt at the computation of the center index. The latter problem can probably be remedied by using the fringe patterns of the collinear HeNe to make accurate thickness measurements[35]. The cloudiness may be corrected by heating the tissue to body temperature.

Similar problems existed during measurement of the lens, with one additional difficulty. The power of the lens caused an intermediate focus between the mirror and the spectrometer and hence the beam travelling through the sample arm was of different width than the beam in the reference arm. No interferograms were obtained for lens.

The refractive index of aqueous humor and vitreous humor measured with white-light interferometry yielded very reliable results for the differential index. The difference between the present measurements and past measurements was usually not greater than 0.0003. However, as the distance increased between the measured position and the stationary phase position, the technique became increasingly inaccurate. This was especially true in the near-infrared when the lower resolution grating was used. In this case, as the fringes became closer to each other, the measured phase became inaccurate because there were merely one or two points between peaks.

Although very accurate results were achieved for the differential index, locating the center index to the same accuracy was not possible. This was due to a number of factors. Although the stage resolution was 100 nm,

this was thought to be sufficient for the measurement. With this resolution, the measurement should have been accurate to 0.0001. That is not to say the stages did not introduce any error in the measurement. Several times in the course of the experiment, the stages were moved several millimeters. Although the resolution when making small steps may have been 100 nm, when making large steps the drift was as large as several microns when returning back to the original position. Another source of error was in locating the stationary phase position without the sample present. The stationary phase position frequently covered more than the entire image and thus was difficult to locate. There is also an error introduced into the measurement when the chosen center peak did not fall exactly on the 256^{th} pixel. This error will be corrected in future updates to the analysis software. Despite the large problems in measurement of the center index, for all the curves presented in this paper, the average error in the measurement of the center index was 0.0036.

From the preliminary work outlined in this paper, there are indications for future work. First, a method for measurement of the index of lens and cornea must be accomplished. Second, ocular components from different species will be measured. Third, the temperature dependence of the refractive index of ocular media will be measured. This will enable slight corrections to be made to past data measured at room temperature, as well as provide insight into the effect of temperature on optical properties. This last finding may have implications in thermally mediated tissue damage. Finally, the ocular tissue will be measured over several days or weeks, to determine how the tissue changes once it is removed from the eye. Since ocular tissue is made up of mostly water, it is not thought that a significant change will be found. However, the study may provide important information for surgical procedures where eye parts from donors are transplanted to living recipients.

7. CONCLUSION

The accurate measurement of the refractive index of fused silica, water, bovine aqueous humor, and bovine vitreous humor were made by white-light interferometry. The maximum error in differential index for fused silica was 0.0003. For the other media, there were too few reference points to calculate the maximum error (there was less than or equal to one reference point for each curve, that point used to set the curve). The measurement of the center index with stage positions had an average accuracy of 0.0036. The values of the index for the four media from 400 nm to 800 nm are listed in the appendix. Use of this technique has the potential to fill in large gaps in the literature and lead to a more accurate predication of chromatic aberration in the eye, as well as more accurate models of light propagation through the eye.

8. ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research. The author would like to thank the assistance of B. A. Rockwell, D. J. Stolarski, R. J. Thomas, and B. Eilert.

REFERENCES

[1]C. Sáinz, J. Calatroni, and G. Tribillon, "Refractometry of liquid samples with spectrally resolved white-light interferometry", *Meas. Sci. Technol.*, **1**, pp. 356-361 (1990).

[2]A.L. Guerrero, C. Sáinz, H. Perrin, R. Castell, and J. Calatroni, "Refractive index distribution measurements by means of spectrally-resolved white-light interferometry", *Opt. Las. Technol.*, **24**, pp. 333-339 (1992).

[3]S. Diddams and J.-C. Diels, "Dispersion measurements with white-light interferometry", *J. Opt. Soc. Am. B*, **13**, pp. 1120-1129 (1996).

[4]D. J. Stolarski, R. J. Thomas, G. D. Noojin, D. J. Payne, and B. A. Rockwell, "White-light interferometric measurements of aqueous media dispersive properties," in *Laser-Tissue Interaction VIII*, S.L. Jacques, Ed., Proc. SPIE, **2975**, pp. 155-162 (1997).

[5]E. Moreels, C. de Greef, and R. Finsy, "Laser light refractometer," *Appl. Opt.*, **23**, pp. 3010-3013 (1984).

[6]K. Kuhler, E. L. Dereniak, and M. Buchanan, "Measurement of the refractive index of solutions at high concentrations," *Appl. Opt.*, **30**, pp. 1711-1714 (1991).

[7]S. Nemoto, "Measurement of the refractive index of liquid using laser beam displacement," *Appl. Opt.*, **31**, pp. 6690-6694 (1992).

[8]H. M. Dobbins and E. R. Peck, "Change of refractive index of water as a function of temperature," *J. Opt. Soc. Am.*, **63**, pp. 318-320 (1973).

[9]G. Abbate, A. Attanasio, U. Bernini, E. Ragozzino, and R. Somma, "The direct determination of the temperature dependence of the refractive index of liquids and solids," *J. Phys. D*, **9**, pp. 1945-1951 (1976).

[10]W. Lu and W. M. Worek, "Two-wavelength interferometric technique for measuring the refractive index of salt-water solutions," *Appl. Opt.*, **32**, pp. 3992-4002 (1993).

[11]D. Stolarski, G. Noojin, R. Thomas, and B. Rockwell, " Interferometric measurement of index of refraction as a function of wavelength in ocular media", in *Laser-Tissue Interaction VII*, S.L. Jacques, Ed., Proc. SPIE, **2681**, pp. 420-426 (1996).

[12]B. Richerzhagen, "Interferometer for measuring the absolute refractive index of liquid water as a function of temperature at 1.064 μm," *Appl. Opt.*, **35**, pp. 1650-1653 (1996).

[13]Ph. Marteau, G. Montixi, J. Obriot, T. K. Bose, and J. M. St. Arnaud, "Simple method for the accurate determination of the refractive index of liquids in the infrared", in *Infrared Technology XVI,* Proc. SPIE, **1341**, pp. 275-278 (1990).

[14]A. Suhadolnik, A. Babnik, and J. Mozina, "Refractive index measurement with optical fiber Mach-Zender interferometer", Proc. SPIE, **1796,** 364, (1992).

[15]B. J. Pernick, "Nonlinear regression analysis for the Sellmeier dispersion equation of CdS," *Appl. Opt.*, **22**, pp. 1133-1134 (1983).

[16]R. H. H. Kröger, "Methods to estimate dispersion in vertebrate ocular media", *J. Opt. Soc. Am. A*, **9**, pp. 1486-1490 (1992).

[17]C. K. Carniglia, K. N. Schrader, P. A. O'Connell, and S. R. Tuenge, "Refractive index determination using an orthogonalized dispersion equation," *Appl. Opt.*, **28**, pp. 2902-2906 (1989).

[18]G. Westheimer, "Optical properties of vertebrae eyes," in *Physiology of Photoreceptor Organs, Vol. VII/2 Handbook of Sensory Physiology*, M. G. F. Fuortes, Ed., Springer-Verlag, N. Y., p. 453 (1972).

[19]Y. Le Grand and S. G. El Hage, *Physiological Optics*, Springer-Verlag, N. Y., pp. 65-66 (1980).

[20]A. Hughes, "The topography of vision in mammals of contrasting life style: comparative optics and retinal organisation," in *The Visual System in Vertebates*, F. Crescitelli, Ed., Springer-Verlag, N. Y., pp. 652-653 (1977).

[21]A. Hughes, "A useful table of reduced schematic eyes for vertebrates which includes computed longitudinal chromatic aberrations," *Vis. Res.*, **19**, pp. 1273-1275 (1979)

[22]C. J. Murphy and H. C. Howland, "The optics of comparative ophthalmoscopy," *Vision Res.*, **27**, pp. 599-607 (1987).

[23]J. G. Sivak and T. Mandelman, "Chromatic dispersion of the ocular media," *Vision Res.*, **22**, pp. 997-1003 (1982).

[24]T . Mandelman, and J. G. Sivak, "Longitudinal chromatic aberration of the vertebrate eye", *Vision Res.*, **23**, pp. 1555-1559 (1983).

[25]A. Chaudhuri, P. E. Hallett, and J. A. Parker, "Aspheric curvatures, refractive indices, and chromatic aberration for the rat eye," *Vision Res.*, **23**, 1351-1361 (1983).

[26]S. Nakao and S. Fujimoto, "Model of the refractive-index distribution in the rabbit crystalline lens," *J. Opt. Soc. Am.*, **58**, pp. 1125-1130 (1968).

[27]R. H. H. Kröger and M. C. W. Campbell, "Dispersion and longitudinal chromatic aberration of the crystalline lens of the African cichlid fish *Haplochromis burtoni*," *J. Opt. Soc. Am. A*, **13**, pp. 2341-2347 (1996).

[28]P. Simonet and M. C. W. Campbell, "The optical tranverse chromatic aberration on the fovea of the human eye," *Vision Res.*, **30**, pp. 187-206 (1990).

[29]B. Gilmartin and R. E. Hogan, "The magnitude of the longitudinal chromatic aberration of the human eye between 458 and 633-nm", *Vision Res.*, **11**, pp. 1747-1753, (1985).

[30]C. Sáinz, P. Jourdain, R. Escalona, and J. Calatroni, "Real time interferometric measurements of dispersion curves", *Optics Communications*, **111**, pp. 632-641 (1994).

[31]Chromex Technical Note 3, "Grating selection parameters," November, 1993.

[32]W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, N. Y., p. 63 (1990).

[33]Melles Griot Catalog "Synthetic Fused Silica," Optics Guide 1997-1998, pp. A4.11-A4.13.

[34]B. A. Rockwell, D. X. Hammer, P. K. Kennedy, R. Amnotte, B. Eilert, J. J. Druessel, D. J. Payne, S. Phillips, D. J. Stolarski, G. D. Noojin, R. J. Thomas, and C. P. Cain, "Retinal laser spot size with wavelength", in *Laser-Tissue Interaction VIII*, S.L. Jacques, Ed., Proc. SPIE, **2975**, pp. 148-154 (1997).

[35]L. Rovati, U. Minoni, and F. Docchio, "Dispersive white light combined with a frequency –modulated continuous-wave interferometer for high-resolution absolute measurement of distance," *Opt. Lett.*, **22**, pp.850-852 (1997).

APPENDIX

Table 2: Refractive index data from 415 nm to 750 nm. Data in italics indicate discrepancy of more than 0.0002 between two measurements. Individual curves are listed together.

λ	WATER		FUSED SILICA		BOVINE VITREOUS		BOVINE AQUEOUS	
415					1.3407		1.3414	
420					1.3403		1.3411	
425			1.4675		1.3400		1.3408	
430	1.3400		1.4671		1.3397		1.3405	
435	1.3397		1.4667		1.3394		1.3402	
440	1.3393		1.4663		1.3391		1.3399	
445	1.3390		1.4660		1.3388		1.3396	
450	1.3387		1.4656		*1.3386*	*1.3389*	1.3394	
455	1.3384		1.4653		*1.3383*	*1.3386*	1.3392	
460	1.3381		1.4649		*1.3381*	*1.3384*	1.3389	
465	*1.3378*	*1.3372*	1.4645			1.3381	1.3386	
470		1.3369	1.4643			1.3379	1.3384	
475		1.3366	1.4640			1.3377	1.3382	
480		1.3364	1.4637			1.3374	1.3379	
485		1.3361	1.4634			1.3372	1.3377	
490		1.3359	1.4627			1.3370	1.3375	
495		1.3357	1.4624			1.3368	1.3374	
500		1.3354	1.4622			1.3366	1.3372	
505		1.3352	1.4620			1.3364	1.3371	
510		1.3350	1.4619			1.3362	1.3369	
515		1.3348	1.4615			1.3360	1.3368	
520		1.3345	1.4613			1.3359	1.3367	
525		1.3344	1.4611			1.3357	1.3365	
530		1.3342	1.4610			1.3355	1.3364	
535		1.3340	*1.4608*	*1.4605*			*1.3363*	*1.3356*
540		1.3338	*1.4606*	*1.4603*				1.3354
545		1.3337		1.4601				1.3352
550		1.3335		1.4599		1.3351		1.3351
555		1.3333		1.4598		1.3350		1.3349
560		1.3332		1.4596		1.3348		1.3347
565		1.3330		1.4594		1.3346		1.3346
570		1.3324		1.4592		1.3344		1.3344
575		1.3324		1.4591		1.3342		1.3343
580		1.3322		1.4589		1.3340		1.3342
585		1.3321		1.4586		1.3339		1.3340
590		1.3320		1.4584		1.3337		1.3339
595		1.3319		1.4581			*1.3335*	*1.3338*
600		1.3318		1.4579			*1.3333*	*1.3337*
605		1.3317		1.4577			*1.3332*	*1.3335*
610	*1.3309*	*1.3317*		1.4575			*1.3331*	*1.3334*
615	*1.3308*	*1.3316*		1.4572			*1.3330*	*1.3333*
620	*1.3308*	*1.3315*		1.4571			*1.3329*	*1.3332*
625	*1.3307*	*1.3314*		1.4570			*1.3328*	*1.3331*
630	*1.3307*	*1.3313*		1.4569			*1.3327*	*1.3330*
635	*1.3306*	*1.3312*		1.4568			*1.3326*	*1.3329*
640	*1.3305*	*1.3312*		1.4567		1.3321	*1.3325*	*1.3328*
645	*1.3305*	*1.3311*		1.4566		1.3321	*1.3324*	*1.3327*
650	*1.3304*	*1.3310*		1.4565		1.3320	1.3323	
655	*1.3304*	*1.3310*		1.4564		1.3319	1.3322	
660	1.3303			1.4563		1.3319	1.3321	
665	1.3303			1.4562		1.3318	1.3320	
670	1.3302			1.4561		1.3317	1.3319	
675	1.3302			1.4560		1.3317	1.3318	
680	1.3302			1.4560		1.3316	1.3318	
685	1.3301			1.4559		1.3316	1.3317	
690	1.3301			1.4558		1.3315	1.3316	
695	1.3301			1.4557		1.3315	1.3315	
700	1.3300			1.4557		1.3314	1.3314	
705	1.3300					1.3314	1.3314	
710						1.3313	1.3313	
715						1.3313	1.3312	
720						1.3312	1.3312	
725						1.3312	1.3311	
730						1.3311	1.3310	
735						1.3311	1.3310	
740						1.3311	1.3309	
745						1.3310	1.3309	
750						1.3310	1.3308	

GENDER EFFECTS IN WAYFINDING STRATEGY: IMPLICATIONS FOR TEAM AND INDIVIDUAL TRAINING

Catherine R. Harrison
Ph.D. Student
Department of Psychology

University of Illinois at Urbana-Champaign
Beckman Institute, HPP Group
405 N. Mathews
Urbana, IL 61801

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong laboratory

August 1997

GENDER EFFECTS IN WAYFINDING STRATEGY: IMPLICATIONS FOR TEAM AND INDIVIDUAL TRAINING

Catherine R. Harrison
Ph.D. Student
Department of Psychology
University of Illinois at Urbana-Champaign

ABSTRACT

Data from two studies are presented which suggest that individuals choose differing strategies for wayfinding, based upon their strengths and preferred cognitive styles. While one strategy does not necessarily convey a benefit for real-world navigation, the literature indicates that individuals who use strategies that emphasize orientation rather than object location tend also to have an advantage in certain spatially-related tasks that demand dynamic re-orientation, such as Unmanned Arial Vehicle (UAV)operation.

Because of the potential consequences of strategy choice, a study is proposed that will test whether a training manipulation can affect a particular individual's choice of strategy for wayfinding. The outcome of the proposed study will have important implications for the training of individuals to perform such spatially demanding tasks as the operation of UAVs.

GENDER EFFECTS IN WAYFINDING STRATEGY: IMPLICATIONS FOR TEAM AND INDIVIDUAL TRAINING

Catherine R. Harrison

INTRODUCTION

My research interest is in identifying different strategies that individuals may adopt in order to solve spatial problems. One example of a spatial problem is wayfinding, which involves finding one's way from a starting position to a goal position. This is something we all do every day. Individual differences in wayfinding ability have long been studied in the scientific literature.

Thorndyke(Thorndyke & Hayes, 1982) elaborated a theory of human acquisition of spatial knowledge consisting of three stages. The first level of mastery of a new space is characterized by acquisition of "landmark knowledge". At this stage the individual orients to highly salient landmarks, and knowledge of the environment consists mainly of visual images of these landmarks. The second level of mastery is characterized by "route knowledge", in which a learned series of procedures permits navigation from one landmark to another. These procedures are of the simple, egocentric "start here, turn left, turn right, go forward" type, and are not readily transferable to novel starting points. The ultimate level of spatial mastery is characterized by "survey knowledge", in which the individual possesses an internal cognitive representation or "cognitive map" of an entire space, including salient landmarks. This level of mastery permits flexible navigation given novel starting points and goals within the mastered space.

Thorndyke's conception of spatial learning is well accepted and has formed the foundation of subsequent research on wayfinding and spatial learning throughout the 80s and 90s (Wickens, 1992). Individual differences researchers, guided by this conception of spatial learning, have identified a number of spatial factors, most notably spatial relations (Sr), spatial orientation (So), and visualization (Vz) (Lohman, 1979), which are considered to represent the basic cognitive abilities that support performance of spatial tasks.

There is a problem with the Thorndyke conception of spatial learning, however. Mounting evidence indicates that women are likely to follow a very different process toward forming survey knowledge, one which does not emphasize route learning but instead emphasizes object-location memory (McBurney, Gaulin, Devineni & Adams, 1997) (Harrison et al., 1997) (Seymore, Dou & Juraska, 1996) (Warren, Harrison, Wszalek, Glass & Kramer, 1997). Note that in the Thorndyke conception, an emphasis upon object-location memory would be categorized as landmark knowledge -- the most elementary level of spatial knowledge. Women perform poorly with respect to men in the speed of acquisition of routes (Harrison et al., 1997) and in the performance of laboratory mental rotation tasks which seem to correlate well with speed of route learning (many references, including (Linn & Petersen, 1986) (Levy & Heller, 1992) (Hampson & Kimura, 1988) (Hampson, 1990)). What is only beginning to be explored is the notion that women may be employing an alternative spatial learning strategy which follows a longer time course than the male-preferred route learning strategy, but which ultimately leads to more robust and detailed survey knowledge of a space (McBurney et al., 1997) (Seymore et al., 1996). Because measures of object-location memory have been conspicuously absent from most spatial learning research, these measures have also been omitted from factor analytic

studies, seemingly confirming the notion that females are at a significant disadvantage in the spatial domain.

Nature versus Nurture in Spatial Ability

There is almost certainly a strong biological influence upon spatial strategy choice and performance in various spatial tasks. For one thing, spatial task performance has been shown to fluctuate with levels of the human sex hormones testosterone and estrogen (Hampson & Kimura, 1988) (Hampson, 1990) (Chiarello, McMahon & Schaefer, 1989). For another thing, rats show the same patterns of gender differences in spatial task performance and learning as do humans (Greenough & Green, 1981) (Juraska, 1991) (Seymore et al., 1996) (Juraska, Henderson & Muller, 1984).

Biological influence notwithstanding, there is strong evidence that experience also plays a strong role in the development of spatial abilities. For example, rats of both genders greatly improve their maze-learning performance when reared in enriched environments that provide ample opportunities for exploration and problem-solving (Greenough & Green, 1981) (Juraska et al., 1984). Law, Pelligrino, and Hunt (Law, Pelligrino & Hunt, 1993) have found that video game participation predicts 50% of the gender-related variance in judging relative velocity of moving targets. Hart and Battiste (1991) report that video game training enhances performance in flight school and that prior video game playing experience correlates with successful performance in flight school. Certainly it may be that only spatially skilled individuals choose to play video games for pleasure; but Subrahmanyam and Greenfield have found that when children of both genders were given intensive practice in the video game Marble Madness, the gender difference was significantly reduced in subsequent

performance on a spatial skills battery (Subrahmanyam & Greenfield, 1994).

Goals

My primary goal during my AFOSR summer was to characterize the spatial learning strategies of individuals who successfully develop survey knowledge of a space, and to determine whether my existing data support the notion that men and women tend to use differing strategies for wayfinding. Based on the outcome of these analyses, I propose a follow-on study to determine whether training can alter a subject's choice of wayfinding strategy, and in turn alter the kinds of information the subject learns about an environment.

STUDY 1

I developed a virtual reality (VR) computerized maze environment, called the Virtual Reality Wayfinding Maze (VRWM) in order to control the kinds of environmental stimuli available to subjects during wayfinding. The VRWM permits systematic manipulation of navigationally useful stimuli in order to probe the subject's knowledge of the environment during learning and after learning has occurred.

Study 1 consisted of administering the VRWM task and probes to elderly and college-aged subjects, along with a number of other cognitive tasks. The goal of this study was to characterize the pattern of decrements in wayfinding-related abilities that occur with normal aging, and to determine if such patterns support the hypothesis that, in general, females and males use differing strategies for wayfinding.

METHOD

Subjects 87 elderly and 56 college-aged subjects participated in Study 1. Elderly subjects were recruited by newspaper advertisement in the city paper of Urbana-Champaign, Illinois, and were paid $7 per

hour for their participation in this study. There were 30 male and 57 female elderly subjects, ranging in age from 55 (the minimum permissible age for participation in the study) to 75 years. 56 college-aged subjects were recruited by advertisement on the campus of the University of Illinois at Urbana-Champaign, and were paid $5 per hour. They ranged in age from 17 to 31 years. There were 38 female and 18 male college-aged subjects.

Elderly subjects were given a screening physiological exam, including tests of cardiac fitness suitable to participate in an exercise program safely, and subjects who did not pass the physical exam were excluded from these experiments.

Apparatus Computerized tasks were administered on an Apple PowerMac computer with 15″ high-resolution video graphics monitors, or on a 386-based PC with a 14″ high-resolution video graphics monitor (see Appendix to determine which was used in each task). For all tasks, subjects were seated in a small, comfortably lit room and responded with keypresses on a standard Macintosh keyboard. Subjects were alone in the testing room for all tests except the spatial navigation maze learning task, during which an experimenter was seated behind the subject in order to record data.

Procedure After practicing maneuvers in a video-game like Virtual Reality environment, subjects were placed in a VRWM maze comprised of 5 square rooms. Each room had 4 doors, only one of which opened to the next room. The doors were opaque so that the subject could not see which door was correct. The rooms were visibly numbered, and within each room each door had a different color, so that subjects could use door colors as a mnemonic cue in subsequent runs. Subjects started in room 1 and were told to find their way to room 5. They were instructed that they would have 6 trials to navigate the same maze, that their time was being recorded, and that the best way to minimize their

time on subsequent runs was to minimize their errors. Although total number of errors and total run time was recorded for each run, the number of errors on run 6, the last run, was the primary dependent measure.

After the 6th run, subjects were given an additional run, called the Transfer Trial, in a VRWM that was identical in every way except that the door color cues had been removed. This trial penalized reliance upon a memorized sequence of door colors, and rewarded a route-based strategy independent of door color.

After the Transfer Trial, subjects were given gridded paper and asked to draw a map of the maze. This task was called the Survey Map Task.

RESULTS

An important result emerged from Study 1. Although there was a significant effect of age on Run 6 Errors, there was neither gender effect nor age by gender interaction. In other words, males and females within each age group were equally skilled (or impaired) in learning to navigate the maze over the course of 6 learning trials.

However, when the Transfer Trial demanded that the subject be able to navigate the same space without the use of the object-location cues (the colored doorframes), using only route knowledge, a significant gender effect emerged, $F(1,139)=4.157$, $p=.04$. It appears that females of both age groups relied more than males upon object-location information (landmarks) rather than upon procedural knowledge of a route through the maze. This result supports the hypothesis that females tend to use an object-location strategy for wayfinding, while males tend to rely upon a route-learning strategy. Table 1 summarizes the data from Study 1.

Table 1. Study 1 means by age group and gender

	Old				Young			
	Male		Female		Male		Female	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Run 6 Errors [A]	3.13	3.7	3.61	3.6	.055	.23	.184	.56
Transfer Trial Errors [A,B,C]	3.13	2.6	5.82	4.6	1.06	1.6	1.05	3.2
Survey Map Errors [A,B]	1.60	1.0	2.21	.89	.222	.55	.447	.76

[A] Age effect, alpha=.05
[B] Gender effect, alpha=.05
[C] Age*Gender interaction, alpha=.05

There was a significant age by gender interaction in Transfer Trial errors, the only age by gender interaction found in these data. This interaction may suggest that younger females, who have been able to form a better survey representation of the space than the older females (as demonstrated by the large age effect in Survey Map Errors, $F(1,138)=101.4$, $p<.001$), are able to make use of their cognitive maps to navigate in the absence of the object-location color cues. If this were the case, the elderly females with their weaker survey representations would not be able to make use of this additional information.

As stated above, both male and female elderly were significantly impaired with respect to college aged subjects in their ability to form a survey representation of the space, as indicated by performance in the map drawing task, $F(1,138)=101.4$, $p<.001$. However, there was also a highly significant gender effect in this task, with males making fewer errors in drawing the Survey Map, $F(1,138)=7.226$, $p=.008$. This is the kind of result that in the past has been taken as evidence of inferior female spatial ability (Thorndyke & Hayes, 1982). However, this result would also be consistent with the findings of Seymore et al (1996) and Juraska et al (1984) that females have a longer time course of attaining survey knowledge, because their strategy demands that they acquire more detailed information about the space in the course of their learning.

This, in turn, makes their survey knowledge, once attained, more robust in the face of environmental changes.

STUDY 2

Study 2 was designed to further test the existence of the two strategies by demonstrating two distinct patterns of performance on a variety of probe tasks that test knowledge of landmarks versus knowledge of routes from the VRWM.

METHOD

Subjects 24 college-aged subjects participated in Study 2. Subjects were recruited by advertisement on the campus of the University of Illinois at Urbana-Champaign, and were paid $5 per hour. They ranged in age from 17 to 31 years. There were 16 female and 8 male college-aged subjects. Half of the females were tested on their first day of menstruation, and the other half were tested two days prior to ovulation, in order to control the effects of estrogen fluctuation on female performance in the spatial tasks (Hampson & Kimura, 1988) (Hampson, 1990)).

Apparatus Apparati used were the same as for Study 1.

Procedure Data were collected over a single three-hour session per subject. Table 2 lists the cognitive paradigms used in sequential order. The VRWM paradigm was similar to that used in Study 1, except that a more complex 8-room maze was used, and instead of the simple door-color object-location cues, each room in the maze was furnished with a unique grouping of realistic furniture. In addition to the Transfer Task and Survey Map probes, two additional probe tasks were added to test the subjects' knowledge of the maze. The first, Maze Landmark Use, presented the subject with novel views of the furniture groupings found in the maze. These views were presented out of the

sequence in which the subject had encountered them within the maze. The subject's task was to make use of the directional properties of the furniture grouping to indicate which direction (forward, left, right, or back) would lead to the next room in the maze. The second new probe task was a simple test of Object Recognition for furniture items that occurred in the maze.

Table 2. Cognitive Experimental Paradigms Used in Study 2

Name of Paradigm	Description
WMS-R Story with immediate recall	A standard neuropsychological test of verbal working memory
WMS-R Visual Reproduction with immediate recall	A standard neuropsychological test of visual working memory
Verbal Fluency	A standard neuropsychological test of the fluency of word production
WMS-R (or WAIS) Digit Span	A standard neuropsychological test of working memory span
WMS-R Story 30-minute delayed recall	A standard neuropsychological test of verbal long-term memory
WMS-R Visual Reproduction 30-minute delayed recall	A standard neuropsychological test of visual long-term memory
Virtual Reality Wayfinding Maze (VRWM)	Tests ability to learn to wayfind in a novel environment. 6 learning trials with landmark objects (furniture groupings)
Maze Transfer Trial	1 transfer trial, landmark objects (furniture groupings) removed
Map Drawing Probe Task	Test of survey knowledge of Wayfinding Maze; subject must draw a bird's eye view map of arrangement of rooms in the maze
Maze Landmark Use	Relational Probe Task (8 trials, 1 per room, order random, no door colors, only 1 chance to pick correct door)
Maze-object recognition	Probe of recognition of objects from maze (yes-no recognition, 16 trials, 8 targets and 8 distracters)
Embedded Figures	Tests ability to detect hidden object borders within complex figures
Mental Rotation	Tests ability to mentally rotate a visualization of a geometric form

The goal of the experiment was to characterize the cognitive strengths and wayfinding strategies of successful navigators. For this reason, an exclusion criterion was based upon satisfactory performance in the learning trials. Subjects who committed more than 9 errors over

runs 5 and 6 combined were excluded from further analysis. This exclusion criterion eliminated 4 females and 2 males, equal proportions of each group.

RESULTS

There were significant effects of gender favoring males in performance on the mental rotation task, $F(1,16)=22.1$, $p<.001$, and the digit span task, $F(1,16)=7.3$, $p=.016$. There was a significant effect of gender favoring females in performance on the verbal fluency task, $F(1,16)=5.36$, $p=.03$. There were no additional significant gender differences in these data. Table 3 presents means for the tasks used in Study 2.

Table 3. Means by gender in Study 2, with significance indicated at alpha=.05

Name of Paradigm	Males		Females		p=
	Mean	SD	Mean	SD	
Story immediate recall	32.3	2.9	29.9	9.3	.55
Visual Reproduction immediate recall	40.5	1.2	40.8	.58	.43
Verbal Fluency	**29.7**	**21.2**	**57.9**	**25.7**	*** .03**
Digit Span	**49.7**	**26.2**	**23.7**	**15.0**	*** .02**
Story delayed recall	25.9	5.7	27.9	10.1	.66
Visual Reproduction delayed recall	28.8	3.5	40.6	.79	.11
Maze Run 6 (VRWM)	.50	.55	1.25	1.48	.25
Maze Transfer Trial	1.17	1.6	2.42	2.3	.26
Survey Map Probe Task	4.67	1.4	4.75	1.7	.92
Maze Landmark Use	2.17	.41	2.00	.95	.69
Maze-object recognition	10.3	2.1	10.4	1.5	.97
Embedded Figures	294.3	112	300.5	122	.92
Mental Rotation	**14.7**	**2.0**	**7.83**	**3.2**	*** <.01**

These gender effects were consistent with previously reported gender differences in the literature (Linn & Petersen, 1986) (Levy & Heller, 1992) (Hampson & Kimura, 1988) (Hampson, 1990)). Perhaps equally important is what was not found. While a comparison of means in the Maze Transfer trial shows a trend toward a gender difference that might become significant with sufficient sample size, it seems clear that in this VRWM there is no gender difference in the ability to form a survey representation of the space, as indicated by the means in the Survey Map task. This is in direct contradiction with prior findings of female inferiority in tests of acquiring a survey representation of a space (Thorndyke & Hayes, 1982). Prior studies have tended to use spatial learning spaces that were stimulus-deprived in comparison to the stimulus-rich VRWM with object landmarks. Perhaps the discrepancy between this study and the findings of Thorndyke and others with regards to gender differences in the formation of Survey Representations is a result of the prior studies providing stimuli needed for successful route-based learning but not for successful object-location-based learning. This question warrants further study and perhaps a meta-analysis of the literature on gender differences in wayfinding. The meta-analysis would focus on the types of environmental cues provided for spatial learning.

In order to further characterize the patterns of performance on the maze-related probe tasks, a Pearson's correlation matrix was calculated. Mental rotation correlates at $r=.762$, $p<.001$, with Maze Transfer Trial performance. In other words, people who have strong mental rotation powers are good at route learning, independent of landmarks. This is consistent with prior demonstrations that men excel at mental rotation and prefer a route-based strategy (Linn & Petersen, 1986).

Results from the Maze Landmark Use task indicate that it was not a well-designed probe for strategy differences. Although it was intended to be a strong test of reliance upon landmarks for navigation, data from this task show a moderately strong (though not significant: $r=.30$, $p=.23$) correlation with mental rotation. In retrospect, it seems obvious that showing a room from a novel orientation and asking the subject to indicate the correct exit direction is a task that can be greatly aided by the ability to mentally rotate a visualization of the room, as much as it also demands an accurate memory of the relative locations of objects in the room. Because mental rotation appears from these data as well as from the literature to be the best predictor of a route-based strategy, this task is hopelessly confounded as a test for wayfinding strategy use.

DISCUSSION

As mentioned in the introduction, women may be employing a spatial learning strategy based on object-location, which ultimately may even lead to more robust and detailed survey knowledge of a space (McBurney et al., 1997) (Seymore et al., 1996). The implications of this would require a rethinking of our current approaches to team and individual selection and training. By failing to recognize the benefits of an object-location-based wayfinding strategy, the Air Force may rely too heavily upon spatial testing batteries which do not include measures of object-location ability. Remember that in Study 2, females and males attained equal mastery of survey knowledge of the space, even though there was a male advantage in route knowledge.

On the other hand, there may be particular training situations in which route-learning or object-location learning are of particular performance value. It would be useful to know if individuals who prefer initially to use one strategy may be trained to use the other.

Proposed Follow-on Research

The proposed research will involve applying the VRWM paradigm to investigate whether the individual differences in selection of strategy are amenable to training manipulations. The answer to this question has immediate relevance to training as well as selection for jobs that include a high spatial component, including UAV operation (2-D to 3-D transformations).

Proposed Methodology

A repeated measures design will be employed to determine if a particular individual's choice of strategy for wayfinding can be altered by training. There will be four sessions of testing per subject, with each session lasting approximately 2 hours. Each of the four sessions will feature a VRWM maze that is novel for that session.

In the first session, the subject will be placed in VRWM that simulates a series of connected rooms containing realistic furnishings and instructed to find the endpoint of the maze by trial and error. As the subject has been warned, six of these learning trials will take place. At the end of the six learning trials, four probe tasks will be administered. Probe 1, the Maze Transfer task requires the subject to re-negotiate the same maze, but with the furnishings removed. High performance on this task indicates knowledge of the maze independent of landmarks (either due to a non-landmark strategy, or to robust learning of the maze regardless of strategy). Probe 2, the Survey Map task, requires the subject to draw a two-dimensional route map of the maze. High performance on this task indicates what has been referred to as "survey knowledge" of the space, the ability to flexibly translate 3-D immersive experience with the maze into a map-like representation. This 3-D to 2-D translation is extremely important in a variety of military tasks, including the operation of UAVs. Probe 3, the Maze Object

Recognition task, requires the subject to perform yes-no recognition of objects, half of which were present in the maze and the other half of which could plausibly have been present but were not. High performance on this task indicates that the subject attended to the landmark objects in the maze (which does not necessarily indicate that landmark objects were employed in a useful wayfinding strategy). Probe 4, Maze Landmark Ordering, presents the subject with pictures of 8 objects, one from each room in the maze, and requires the subject to place the object in the order in which they would be encountered while correctly travelling the maze. Because this task does not require mental rotation as the Maze Landmark task of Study 2 did, it should provide a cleaner test of the subject's knowledge and use of landmarks in a powerful navigational manner, indicative of a landmark strategy.

While a subject's response on any one probe task reveals only a limited amount of information, the subject's constellation of performance across these probes does in fact reveal a strategy choice for all but the highest performing subjects (for whom it is not possible to discern the learning strategy that led to their high performance, as they have mastered all aspects of the tasks). This has been documented in my own prior research (Harrison et al, in progress).

For the second and third sessions of testing, the subject will be assigned to one of five conditions. In conditions 1-4, the subject will perform one (the same one within condition, with each condition receiving a different probe) of the four probe tasks after each of the six learning trials. The effect of this will be to train the subject to attend to a particular kind of information within the maze. Subjects in the fifth condition will not receive any probe tasks between maze learning trials, but will receive all four at the end of each session, just as in session 1. Assignment of subjects to conditions will be counterbalanced by gender and by performance scores during the initial

round of testing, in order to match subjects within groups. Each group will contain 5 male and 5 female subjects, for a total N of 50 for the study.

The final testing session will be a repeat of session 1. Subjects may well deduce that they will be given all four probes again, but that is not a problem with the design in that it will serve only to reduce any false finding of a learning effect, rather than accentuate it. Analysis by individual subject will determine whether the training treatment impacted the subjects' performances on any of the probe tasks, and whether the effect varied by gender, by training manipulation group, and by initial strategy preference. The no-training group will permit us to examine the effect of simple practice on performance on the probe tasks and upon strategy.

REFERENCES

Chiarello, C., McMahon, M., & Schaefer, K. (1989). Visual cerebral lateralization over phases of the menstrual cycle: A preliminary Investigation. Brain and Cognition, 11, 18-36.

Greenough, W. T., & Green, E. J. (1981). Experience and the changing brain. In J. L. McGaugh, J. G. March, & S. B. Kiesler (Eds.), Aging: Biology and Behavior (pp. 159-200). New York: Academic Press.

Hampson, E. (1990). Variations in sex-related cognitive abilities across the menstrual cycle. Brain & Cognition, 14(1), 26-43.

Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. Behavioral Neuroscience, 102(3), 456-459.

Harrison, C., Kramer, A., Hahn, S., Bardell, L., Cohen, N., McCauley, E., Vakil, E., & Chason, J. (1997). A novel virtual reality paradigm find realistic patterns of wayfinding impairment with age. Manuscript in preparation.

Hart, S. G., & Battiste, V. (1991). Transfer of Training: Video games to flight. Paper presented at the American Psychological Association, San Francisco, CA.

Juraska, J. (1991). Sex differences in "cognitive" regions of the rat brain. Psychoneuroendocrinology, 16, 105-119.

Juraska, J., Henderson, C., & Muller, J. (1984). Differential rearing experience, gender, and radial maze performance. Developmental Psychobiology, 17(3), 209-215.

Law, D. J., Pelligrino, J. W., & Hunt, E. B. (1993). Comparing the tortoise and the hare: gender differences and experience in dynamic spatial reasoning tasks. Psychological Science, 4(1), 35-40.

Levy, J., & Heller, W. (1992). Gender differences in human neuropsychological function. In H. Gerall, H. Moltz, & L. L. Ward

(Eds.), Sexual Differentiation (Vol. 11, pp. 245-274). New York: Plenum Press.

Linn, M. C., & Petersen, A. C. (1986). A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement. In J. S. Hyde & M. C. Linn (Eds.), The Psychology of Gender (pp. 67-101). Baltimore, MD: Johns Hopkins University Press.

Lohman, D. F. (1979). Spatial ability: A review and reanalysis of the correlational literature (Tech. Rep. No. 8). Stanford, CA: Stanford University.

McBurney, D. H., Gaulin, S. J. C., Devineni, T., & Adams, C. (1997). Superior spatial memory of women: Stronger evidence for the gathering hypothesis. Evolution and Human Behavior, 18, 165-174.

Seymore, P., Dou, H., & Juraska, J. (1996). Sex differences in radial maze performance: influence of rearing environment and room cues. Psychobiology, 24(1), 33-37.

Subrahmanyam, K., & Greenfield, P.-M. (1994). Effect of video game practice on spatial skills in girls and boys. Special Issue: Effects of interactive entertainment technologies on development. Journal of Applied Developmental Psychology, 15(1), 13-32.

Thorndyke, P.-W., & Hayes, R.-B. (1982). Differences in spatial knowledge acquired from maps and navigation. Cognitive Psychology, 14(4), 560-589.

Warren, S., Harrison, C., Wszalek, T., Glass, A., & Kramer, A. (1997). Spatial task performance fluctuates with estrogen levels. Manuscript in preparation.

Wickens, C. D. (1992). Spatial perception and cognition and the display of spatial information, Engineering Psychology and Human Performance (Second ed., pp. 116-166). New York: HarperCollins Publishers, Inc.

EXAMINATION OF AN ORGANIZATIONAL CLIMATE MEASURE AND THE RELATIONSHIP WITH GRIEVANCES AND TRUNOVER

Laura J. Hott
Graduate Student
Department of Psychology

Wright State University
335 Fawcett Hall
Dayton, OH 45345

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

EXAMINATION OF AN ORGANIZATIONAL CLIMATE MEASURE AND THE RELATIONSHIP WITH GRIEVANCES AND TURNOVER

Laura J. Hott
Graduate Student
Department of Psychology
Wright State University

Abstract

An organizational climate survey given to both military and civilian employees at a Midwestern Air Force Base was examined to determine the stability of the measure and the longitudinal effects that climate has on grievance and turnover rates at the organizational level. The climate scales were relatively stable across the years examined. Limited relationships were found between the climate scales and either grievances or turnover. The multidimensionality of organizational climate became evident. In particular, three scales demonstrated a number of significant or marginally significant correlations. The three scales were determined to compose a leadership dimension of climate, and the relation of the leadership dimension with grievance and turnover rates was then examined. A number of limitations and suggestions for future research are discussed.

EXAMINATION OF AN ORGANIZATIONAL CLIMATE MEASURE AND THE RELATIONSHIP WITH GRIEVANCES AND TURNOVER

Laura J. Hott

Introduction

Research has examined the effects that an organization's climate has on outcome variables such as measures of performance, job satisfaction, and employee turnover (e.g., Muttar, 1985; Proctor, Lassiter, & Soyars, 1976; Tyagi & Wotruba, 1993). The purpose of the present study was to examine a measure of organizational climate and to investigate the relations that climate has with grievances and turnover.

Organizational Climate

A number of definitions of organizational climate have been offered, and no single definition is consistently used within the literature. Schneider and Snyder (1975), for example, state that organizational climate is people's perceptions of what the organization is, that these perceptions are based on the events and conditions that exist or take place in the work setting, and that the perceptions are descriptive of the conditions and events. According to Reichers and Schneider (1990), "climate is shared perceptions of organizational policies, practices, and procedures, both formal and informal" (p. 22). Brower (1986) reiterates that climate is a collective perception of individuals about the work environment but adds that the organizational climate affects individual's behavior. The company that developed the climate survey used in this study defined organizational climate as "peoples' perceptions about their organization and department" (Somers, Rouiller, & Hsiao, 1989, p. i). Finally, Cooke and Rousseau (1988) state that "climate, rather than culture, reflects perceptions of organizational structures and how it *feels* to be a member of the organization" (pp. 250-251).

A related concept to organizational climate is organizational culture. Singh and Das (1978) define organizational culture as "a set of attributes and mechanisms that organizations adopt and utilize to deal with their members and environment" (p. 511). Cooke and Rousseau (1988) state that organizational culture is "the shared beliefs and values guiding the thinking and behavioral styles of members" (p. 245).

Reichers and Schneider (1990) discuss four similarities that exist between climate and culture. First, both climate and culture involve the manners by which organizational members make sense of their environment. Second, both are learned, for example, through the socialization process of new members. Third, both are monolithic and multidimensional constructs simultaneously. Finally, both climate and culture attempt to identify the environment that affects the behavior of people within organizations. Differences that have been noted between the two constructs are that culture is more global (e.g., Muttar, 1984; Reichers, & Schneider) and that climate is a manifestation of culture, which exists at a higher level of abstraction (Reichers, & Schneider).

Organizational culture is included in this discussion of climate both because of the relatedness of the two concepts and also due to the scarcity of literature examining the relation of climate and turnover. Thus, additional research will be presented that provides information about the relation of culture and turnover. The reader is

reminded that the concept of interest in the present study is organizational climate, which is defined as the descriptive perceptions that employees have of what occurs or exists in the organization (Schneider & Snyder, 1975). Next, leadership will be discussed as one dimension of climate, which will be followed by a review of the literature pertaining to climate's or culture's relation with turnover.

Leadership dimension of climate. Leadership or supervisor style is one dimension of climate that is included in many measures or conceptualizations of organizational climate. For example, Proctor, Lassiter, and Soyars (1976) surveyed U.S. naval officers using five dimensions, which included upward influence-downward involvement, mutual support, and encouragement of initiatives. These three dimensions all include the subordinates' perceptions of their superior, such as having confidence and trust in superior or the superior demonstrating belief in subordinates' manner of behaving. Tyagi and Wotruba (1993) modified Jones, James, Bruni, Homick, and Sells' (1977) measure of organizational climate and included dimensions assessing leader goal emphasis and work facilitation, leader trust and support, and leader interaction facilitation. Supervisory style was one dimension of climate assessed by Jackofsky and Slocum (1988) in their longitudinal study of climate. Brower (1986) conducted a factor analysis on a climate survey and found that the factor accounting for the most variance was the supervision factor. This factor included items reflecting the employee's perception of the relationship that he/she has with the supervisor and included topics such as the openness of communication, the supervisor's praising of the employee's work, and the employee's overall perception of the supervisor's performance. Finally, in a related culture study, Singh and Das (1978) used a number of culture dimensions to examine the relation between organizational culture and commitment. One culture dimension used was the quality of leadership styles.

Next, a review of the literature demonstrating a relation between organizational climate and turnover will be presented. Again, due to the paucity of literature about this relationship, research examining the relation between organizational culture and turnover will also be presented.

Organizational climate and turnover. There are a few research studies that examine the relation between organizational climate (or culture) and turnover. Proctor et al. (1976) found that perceived organizational climate was negatively related to turnover in a sample of U.S. Naval officers. In addition, stayers perceived their naval unit's organizational climate more favorably then leavers perceived their unit's climate. Bowers (1983) investigated why 11,500 air traffic controllers would strike, which goes against federal law and involves quitting high paying jobs. His conclusion was that the "individual controllers responded to an organizational climate that they experienced as uncaring, unconcerned for its people, uncommunicative, and unreceptive" (p. 16). Jackofsky and Slocum (1988) found that employees' descriptions of their work environment affected intentions to leave along with the amount and type of job satisfaction they possessed. In a study to examine the direction of causality in the relation between intentions to quit and organizational climate, Tyagi and Wotruba (1993) state that quitting intentions were more instrumental in affecting perceptions of organizational climate rather than vice versa. Mobley, Griffeth, Hand, and Meglino (1979) present a model of the employee turnover process that includes climate as an antecedent to job satisfaction, which in turn acts as an antecedent to intentions to quit, which is then a precursor to

turnover behavior or other withdrawal behavior. Finally, Sheridan (1992) found that new employees stayed 14 months longer in organizational cultures emphasizing interpersonal relationship values than in a culture emphasizing work task values. The two dimensions of interpersonal relationship values, team orientation and respect for people, appear to be more leadership oriented than the dimensions encompassing work task values, detail, and stability.

Leadership

As stated earlier, leadership is one dimension of organizational climate. According to Forsyth (1990), "leadership is a reciprocal, transactional, and transformational process in which individuals are permitted to influence and motivate others to promote the attaining of group and individual goals" (p. 216).

A related concept containing leadership and management commitment as themes is Total Quality Management (TQM; Mansir & Schacht, 1989; Steel & Jennings, 1992). TQM is a strategy and approach for the improvement of personal performance and effectiveness throughout an organization and for adjusting and focusing the efforts of individual organizational members to improve the efforts and their related effects (Mansir & Schacht). TQM has been connected with the Baldridge criteria for the Malcom Baldridge National Quality Award, which recognizes U.S. companies that excel in management and achievement quality (Steel & Jennings). One of the Baldridge criteria is leadership, which consists of the leadership system, values, expectations, and public responsibilities of the organization.

Within the literature, the importance of leadership in TQM has also been stressed. For example, Mansir and Schacht (1989) state that effective leadership is essential for effective TQM, including leaders demonstrating commitment to TQM. In addition, according to Deming (1986), leadership is involved throughout the whole process of TQM. Further, Schneider, Brief, and Guzzo (1996) assert that "weak management," such as not honestly and openly interacting with lower level employees, is the prominent reason TQM implementation fails.

Leadership and turnover. Research has investigated relations between leadership style, leader qualities, leader-member exchange (LMX), and employee turnover. Fleishman and Harris (1962) found a curvilinear relationship between a foreman's rating of consideration and both turnover and grievances. Turnovers, along with grievances, were highest in those work groups where the foreman was rated low in consideration. A replication of the Fleishman and Harris' study was conducted by Skinner (1969), who also found a curvilinear relation between consideration and turnover. In an ex post facto study design, Saleh, Lee, and Prien (1965; cited in Porter & Steers, 1973; Mobley, 1982) found that lack of consideration was the second most cited reason for termination, after job content. Ley (1966) found a highly significant correlation ($r = .76$) between turnover and authoritarian ratings of employees' foremen, suggesting that employees are more likely to leave the organization if the foreman or supervisor is authoritarian than if he or she is not authoritarian.

According to Bowers (1983), the decision to strike was easier for air traffic controllers to make due to their dissatisfaction and the poor supervisory leadership. In addition, dissatisfaction, felt burnout, alienation, and lack of job clarity made the decision to stay more difficult for those controllers who stayed on the job. Although striking

behavior is not the same as turnover behavior, Bowers' findings lend support to the relation between leadership and withdrawal behaviors, such as turnover. Krackhardt, McKenna, Porter, and Steers (1981) perceived the supervisor's role as influencing new employees by providing structure, feedback, and rewards to employees, and then these behaviors and the associated subordinates' perceptions may affect the turnover behavior of employees. Krackhardt et al. examined the effect of two interventions on turnover. The two interventions were supervisors informally meeting with each teller and supervisors meeting with all tellers as a group. They found lower turnover in those bank branches participating in the interventions compared to the turnover at the control branches. Hendrix, Hanby, and Zimmerman (1987) tested a model of turnover that included supervisory style and found that a positive supervisory style was significantly related with intent to remain with the organization. Lefkwitz (1971; cited in Steers & Mowday, 1981) found that supervisory style along with other variables (i.e., job expectations, satisfaction, work environment, compensation, and the job itself) influenced turnover.

Mobley (1982) states that a supervisor, by his or her behavior, can aid in managing turnover. Mobley suggests that if the supervisor makes task achievement attainable, for example by removing obstacles and by supplying feedback, employees will be less likely to leave the organization. Further, Mobley proposed investigation of the relationship between LMX and employee turnover. Graen, Liden, and Hoel (1982) examined the ability of individualized leadership (measured by a five-item LMX scale) and of average leadership style (measured by averaging the LMX ratings of all members reporting to a specific supervisor) to predict turnover. The results indicated that individualized LMX, average leadership style, and the deviation between the individualized LMX and the average leadership style were all significantly and negatively related to turnover. In addition, LMX had significant predictive power to make correct predictions about employee turnover, but average leadership style did not. Furthermore, Graen et al. examined the partial correlations between LMX and turnover after controlling for employee attitudes (i.e., satisfaction) and also between employee attitudes and turnover after controlling for LMX. Even after the effects of the employee attitudes were partialed out, LMX remained a significant predictor of turnover. Only one employee attitude, the pay subscale of the Job Descriptive Index, was significantly related to turnover after controlling for LMX. Ferris (19850 conducted a replication of the Graen et al. study. Although he replicated Graen et al.'s findings, the strength of LMX as a predictor of turnover was not as strong as Graen et al. found. Both Graen et al.'s and Ferris' studies found that average leadership style (average of all subordinates under a particular supervisor) was not as strong a predictor of turnover as LMX was (at the dyadic level of supervisor-subordinate). These studies examined turnover at the individual level. The present study, however, investigated the relation between leadership climate and turnover at the organizational level.

Summary

In the present study, climate is defined as the perceptions employees have about the organization, which are descriptive perceptions of what occurs or exists in the organization. Research exists supporting a relation between organizational climate and turnover (e.g., Proctor et al., 1976; Tyagi & Wotruba, 1993). Leadership is one dimension of climate and has been included in a number of climate measures (e.g., Proctor et al.; Jackofsky &

Slocum, 1988). The relation between leadership and turnover and grievances has been demonstrated (e.g., Fleishman & Harris, 1962). A question to be examined is whether the leadership dimension of climate is related to turnover and grievances.

Purpose of Study

The purpose of the present study was to examine the effects of climate on turnover and grievances at the organizational level, specifically focusing on the effects of the leadership dimension of climate on these two outcome variables. In addition, the present study investigated the longitudinal effects of climate on turnover and grievances by exploring the relations between climate responses for one year with grievances and turnover of subsequent years.

Methodology

Participants

The participants came from 34, 32, and 29 functional divisions of an Air Force organization in the Midwest for 1994, 1995, and 1996, respectively. In 1994, 6409 individuals participated with 6042 individuals providing complete data. In 1995, 5439 individuals participated with 4639 individuals providing complete data. In 1996, there were 5061 participants with 4527 individuals providing complete data. Demographic information was available for 65% (25 of the 29 divisions) of the 1996 sample. This information can be considered representative of the three sample years. In 1996, 63.6% of the unit's employees were civilian personnel versus 36.4% being military personnel. Thirty-seven percent of the civilians were female, and 21% of the military personnel were female. Almost 14% of civilians were age 34 or younger, nearly 15% were between the ages of 35 and 39, about 16% were 40 to 44 years old, around 20% were in the age range of 45 to 49, and about 35% were 50 years or older. For military personnel, the majority (65.5%) were age 34 or younger, about 19% were between the ages of 35 and 39, around 11% were in the age range of 40 to 44, about 5% were between 45 and 49 years old, and only about 1% were age 50 or older. The status of the respondents (civilian vs. military) was requested on the 1995 and 1996 surveys.

Measures

Climate. Organizational climate was assessed using the Air Force Cultural Survey. Although this survey is labeled a culture survey in the title, after examining the items, the researchers determined that the survey was assessing climate. The Air Force Cultural Survey consists of 58 items written around the seven Baldridge criteria and written to assess 20 subscales. Participants responded on a six-point scale from strongly disagree to strongly agree. Participation was voluntary and anonymous. Coefficient alphas ranged from .70 to .91 for these 20 scales (Somers et al., 1989). In the present study, the coefficient alphas obtained ranged from .68 to .93 for the 20 a priori scales.

Turnover and grievances. The number of grievances reported and the number of individuals leaving were obtained from the Human Resources department for the fiscal years of 1994, 1995 and 1996. This data was available only for civilian personnel. These measures are at the organization level, and no information was available

at the individual level. In addition, no information was available to the researchers about whether the turnover was voluntary or involuntary.

Analyses

In order to compare the climate responses with the criterion variable data, the responses for civilian personnel were aggregated to the organizational level. Thus, mean scores for each organization on the climate survey subscales were used in subsequent analyses. Secondly, separate factor analyses were conducted using the individual level data for the three years to determine the number of factors using varimax rotation and the criterion that eigenvalues had to be greater than or equal to one. Third, correlations were computed to determine the stability of the scales across time. Thus, the 1994 scales were correlated with both the 1995 and 1996 scales, and the 1995 scales were correlated with the 1996 scales. The fourth planned analysis was to compute the correlations between the climate scales and grievances and turnover for 1995 and 1996. Finally, exploratory analyses were conducted to examine the relations between the leadership component of the climate survey and turnover and grievances individually.

Results

Factor Analysis

For the 1994 data set, eight factors were obtained that met the criteria mentioned earlier. The eighth factor contained only one item. Seven factors were produced for both 1995 and 1996. The items tended to load together, although the ordering of the factors varied. The seven factors, excluding the single item factor for 1994, corresponded with the seven Baldridge criteria. Factor analyses were also conducted on the 1995 and 1996 civilian only data, with only minor changes in the factor structure for 1995 and no changes in the 1996 factor structure. For both samples, the coefficient alpha estimates were computed for the seven Baldridge scales and ranged from .77 to .97.

Stability of Climate Measure Across Time

Next, the scale scores for the 20 a priori and the seven Baldridge scales were each correlated across years using the complete sample.

Baldridge scales. The correlations of the seven Baldridge scales for 1994 with the scales from 1995 resulted in six of the seven correlations being significant. The correlations ranged from .35 to .70. All seven of the correlations between the 1994 and 1996 scales were significant and ranged from .41 to .75. Six of the seven correlations between the Baldridge scales from 1995 and 1996, ranging from .18 to .79, reached significance.

A priori scales. The correlations between the 20 a priori scales from 1994 with those from 1995 ranged from .35 to .80, with 18 of the 20 correlations being significant at the .05 level. The range of the correlations between 1994 and 1996 on the a priori scales was from .17 to .83, with 12 of the 20 correlations achieving significance. The correlations between the 1995 scales and the 1996 scales ranged from .25 to .94, with 4 of the 20 correlations reaching significance.

Correlations of Climate Scales with Grievances and Turnover

The correlations between the climate scales and grievances and turnover were computed using only the civilian data from 1995 and 1996.

Baldridge scales. Three correlations between the Baldridge scales and turnover were significant or marginally significant: Business Results for 1996, Leadership for 1996, and Process Management scale for 1996 all with turnover for 1995 ($r = -.322$, $r = -.314$, $r = -.367$, $p < .10$, respectively). No correlations between the Baldridge scales and grievances were significant (see Table 1).

A priori scales. A few of the scales had significant or marginally significant correlations with either grievances or turnovers (see Table 2). Three scales (Communications, Leader's Visible Commitment to Goals, and System/Structure for Quality Improvement) had between three or four significant or marginally significant correlations with grievances and turnovers. Five other scales (Attitudes/Morale, Innovation, Involvement, Supervisor's Role in Quality Improvement, and Social Interactions) had one correlation with 1995 turnover that was marginally significant. Based on the correlations, it was determined that the first three scales, which had a number of significant or marginally significant correlations, fit together to form a leadership component. This leadership component was then further analyzed through exploratory analyses by regressing the three scales onto either grievances or turnover for the same year and the subsequent years.

Exploratory Analyses

Regression equations were computed using a subset of the a priori scales as independent variables and using either grievances or turnover as the dependent variable. Three sets of equations were analyzed with grievances or turnover as the dependent variable. First, equations used the scales from 1995 with the dependent variable from 1995 and 1996. Second, equations for the 1996 scales used the 1996 dependent variable values.

Equations for 1995 were computed two ways. First all three scales were entered as independent variables. Second, the equations were computed without System/Structure for Quality Improvement as an independent variable because this scale did not have any significant or marginally significant correlations.

Three scales as independent variables. First, regression equations were computed using grievances as the dependent variable and the three scales identified earlier in the correlational (Communications, Leader's Visible Commitment to Goals, and System/Structure for Quality Improvement) as the independent variables. The three scales were entered simultaneously. No equations predicting either 1995 or 1996 grievances were significant or marginally significant. In the regression equations with turnover as the dependent variable, none of the equations predicting either 1995 or 1996 turnover were significant or marginally significant.

Two 1995 scales as independent variables. Next, equations using the 1995 Communications and Leader's Visible Commitment to Goals scales as independent variables were computed. The first set of equations used grievances as the dependent variable, the neither of the equations were significant. The second set of equations used the two scales as independent variables and turnover as the dependent variable. Again, the equations were not significant in predicting either 1995 or 1996 turnovers.

Discussion

The multidimensional nature of climate is evident from the correlational analysis results. In addition, not all aspects of climate appear to be related to turnover and grievances. In the present study, only the leadership dimension consistently displayed a relationship with either turnover and grievances. Perhaps different climate dimensions are related to other outcome variables such as job satisfaction, absenteeism, or performance.

Relationships between climate scales and either grievances or turnover occurred only when the 20 a priori scales were used but not when the seven Baldridge scales were examined, which may further emphasize the multidimensionality of climate. Perhaps the Baldridge scales are related to different outcome variables, especially considering that the criteria are used for a performance and quality based award. Another possible explanation is that the Baldridge scales are too broad in their content, and, thus, relations are not found.

Both the 20 a priori and the seven Baldridge climate scales were relatively stable over time, as demonstrated by the correlations across the years. Two scales that had low stability correlations were Leader's Involvement and System/Structure for Quality Improvement. These low correlations may be the reason that the 1995 and 1996 Leader's Involvement scales and the 1995 System/Structure for Quality Improvement scale did not produce significant or marginally significant correlations.

Based on the exploratory regression analyses, the leadership component of climate does not allow any prediction of grievances and turnover.

Limitations

There are a number of limitations associated with this study that may have affected the results or that could affect the generalizability of the results. First, because both the turnover and grievance data were only available at only the divisional level, the individual level climate data was aggregated to obtain a mean for each item within each division. This aggregation led to the sample size shrinking from numbers in the thousands to a sample size around 30. Because the smaller sample size may have reduced the statistical power, a significance level of .10 was used when examining the results. An additional limitation related to the sample is that a number of divisions were either dissolved or developed during the course of the study. However, the individuals from the dissolving divisions did not necessarily get placed into the developing divisions on a one-to-one ratio. There was no way to track which divisions individuals moved from or to during the course of the study. Furthermore, there was no way to determine if the same individuals responded from year to year. The survey was anonymous, and no information allowing for this type of determination was requested from participants.

The correlational analyses between the climate scales and the grievance and turnover data was conducted using the mean civilian data within divisions, because the data for the criteria were only available for the civilian personnel. In addition, the status of the respondents was requested on the 1995 and 1996 surveys, but not on the 1994 survey. Thus, these analyses could not examine the effect of the 1994 climate responses on grievances and turnovers for 1994, 1995, and 1996.

As was mentioned earlier, the turnover data only indicated the number of individuals leaving a particular division. No information was available about the proportion of voluntary versus involuntary turnover within each division or overall. Perhaps the dimensions of climate are differentially related to voluntary and involuntary turnover.

A fifth limitation of the study is that the wording was different for many survey items in 1995 and 1996 versus the 1994 survey items. Although the wording change may seem insignificant to some, participants' perceptions of the meaning of the items may have changed, which could affect their responses and thus the results of the study. A related limitation is that additional items, which were not included in these analyses, were added in both 1995 and 1996. Again, the additional items could affect the participants' responses and perceptions of the original 58 items.

There are two additional pieces of information that are important for one to be aware of when examining the results of the present study. First of all, the Air Force Base that was used in the study was unusual because of the high proportion of civilian to military personnel as compared to other bases. Military personnel may view the climate differently than civilians do, or perhaps civilians have an easier time leaving the organization or expressing a grievance than military personnel do. The different perceptions of military and civilian personnel may be a reason for some expected climate scales to not produce significant correlations with turnover or grievances. The organization may treat and view civilian personnel differently than military personnel, which in turn could affect individual's perceptions of the climate of the organization. The second piece of information comes from Mobley (1982), who suggests that when individuals want to quit but are unable to, they may engage in other forms of withdrawal (e.g., absenteeism, apathy, sabotage, poor quality work). Other types of withdrawal behaviors, aside from grievance rates, were not investigated in the present study.

Future Research

There are a number of areas where future research is needed or suggested. First, the multidimensionality of climate needs further exploration. As mentioned earlier, a number of researchers have developed measures with multiple dimensions. Further research is needed to determine which dimensions are most important to the organization.

Relations were found between the components of the leadership dimension of climate and grievances and turnover. Research investigating the relations between other dimensions of climate and both grievances and turnover are necessary. Other dimensions may demonstrate important relations with grievances and turnover compared to the leadership dimension. Also, some aspects of the leadership dimension may be more related to grievances or turnover.

Third, all of the analyses for the present study were conducted at the divisional level due to the nature of the data. Research in the future should explore the relations between climate dimensions, especially the leadership dimension, with grievances and turnover at the individual level. The individual employee's perceptions of climate may have a profound effect on whether or not a grievance is expressed or on if the employee leaves the

organization. In addition, both voluntary and involuntary turnover should be examined with respect to climate dimensions because different processes may be occurring with these two types of turnover. It may be that only voluntary turnover is related to a particular dimension of climate such as leadership, but no relationship exists between that dimension and involuntary turnover.

Other outcome variables, such as performance, and their relation with the dimensions of climate should be examined. It would also be advantageous to investigate other withdrawal behaviors, including absenteeism or poor quality work, to determine if these variables are related to climate dimensions. Once a more accurate representation of what influences the outcomes determined, suggestions for decreasing the frequency of the negative behaviors can be presented.

Finally, the Air Force Base where this study was conducted was unusual due to the proportion of civilians that are employed. Many Air Force Bases have about a three to one ratio of military personnel to civilians. In contrast, the particular base used in the study has roughly a two to one ratio of civilians to military personnel. Future research should examine the relations of climate to turnover and grievances, along with other outcome variables, at other Air Force Bases or other organizations. The results found in the present study may not be generalizable to other Air Force Bases or to non-military organizations.

References

Bowers, D. G. (1983). What would make 11,500 people quit their jobs? Organizational Dynamics, 11(3), 5-19.

Brower, R. I. (1986). An analysis of an industrial corporation's employee climate survey and job satisfaction from 1980 to 1984. Unpublished doctoral dissertation, University of Kansas, Lawrence.

Cooke, R. A., & Rousseau, D. M. (1988). Behavioral norms and expectations: A quantitative approach to the assessment of organizational culture. Group and Organization Studies, 13(3), 245-273.

Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Center for Advanced Engineering Study, MIT Press.

Ferris, G. R. (1985). Role of leadership in the employee withdrawal process: A constructive replication. Journal of Applied Psychology, 70(4), 777-781.

Fleishman, E. A., & Harris, E. F. (1962). Patterns of leadership behavior related to employee grievances and turnover. Personnel Psychology, 15, 43-56.

Forsyth, D. R. (1990). Group dynamics (2nd ed.). Pacific Grove, CA: Brooks/Cole Publishing.

Graen, G. B., Liden, R. C., & Hoel, W. (1982). Role of leadership in the employee withdrawal process. Journal of Applied Psychology, 67(6), 868-872.

Hendrix, W. H., Hanby, G. E., & Zimmerman, B. K. (1987). Correlates of auditors' intent to remain with present organization. Journal of Social Behavior and Personality, 2(1), 79-92.

Jackofsky, E. F., & Slocum, J. W., Jr. (1988). A longitudinal study of climates. Journal of Organizational Behavior, 9, 319-334.

Krackhardt, D., McKenna, J., Porter, L. W., & Steers, R. M. (1981). Supervisory behavior and employee turnover: A field experiment. Academy of Management Journal, 24(2), 249-259.

Ley, R. (1966). Labor turnover as a function of worker differences, work environment, and authoritarianism of foremen. Journal of Applied Psychology, 50(6), 497-500.

Mansir, B. E., & Schacht, N. R. (1989). Total quality management: A guide to implementation. Bethesda, MD: Logistics Management Institute.

Mobley, W. H. (1982). Employee turnover: Causes, consequences, and control. Reading, MA: Addison-Wesley.

Mobley, W., Griffeth, R., Hand, H., & Meglino, R. (1979). A review and conceptual analysis of the employee turnover process. Psychological Bulletin, 86(3), 493-522.

Muttar, K. A. (1984). An investigation of the validity of objective and subjective measures of organizational climate. Unpublished doctoral dissertation, West Virginia University, Morgantown.

Porter, L. W., & Steers, R. M. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. Psychological Bulletin, 80(2), 151-176.

Proctor, J. H., Lassiter, W. E., & Soyars, W. B., III (1976). Prediction of young U.S. naval officer retention. Personnel Psychology, 29, 567-581.

Reichers, A. E., & Schneider, B. (1990). Climate and culture: An evolution of constructs. In B. Schneider (Ed.), Organizational climate and culture (pp. 5-39). San Francisco: Jossey Bass.

Schneider, B., Brief, A. P., & Guzzo, R. A. (1996). Creating a climate and culture for sustainable organizational change. Organizational Dynamics, 24(4), 7-19.

Schneider, B., & Snyder, R. A. (1975). Some relationships between job satisfaction and organizational climate. Journal of Applied Psychology, 60(3), 318-328.

Sheridan, J. E. (1992). Organizational culture and employee retention. Academy of Management Journal, 35(5), 1036-1058.

Singh, P., & Das, G. S. (1978). Organizational culture and its impact on commitment to work. Indian Journal of Industrial Relations, 13(4), 511-524.

Skinner, E. W. (1969). Relationships between leadership behavior patterns and organizational-situational variables. Personnel Psychology, 22, 489-494.

Somers, R. L., Rouiller, J. Z., & Hsiao, D. W. (1989). Development of productivity and quality self-evaluation audits (Technical Report No. 1582-03-89-CR). McLean, VA: General Research Corporation.

Steel, R. P., & Jennings, K. R. (1992). Quality improvement technologies for the 90s: New directions for research and theory. Research in Organizational Change and Development, 6, 1-36.

Steers, R. M., & Mowday, R. T. (1981). Employee turnover and post-decision accommodation processes. Research in Organizational Behavior, 2, 235-281.

Tyagi, P. K., & Wotruba, T. R. (1993). An exploratory study of reverse causality relationships among sales force turnover variables. Journal of the Academy of Marketing Science, 21(2), 143-153.

Table 1

Correlations between Baldridge Scales and Grievances and Turnover for 1995 and 1996

Baldridge Scale	1995 Grievances	1996 Grievances	1995 Turnover	1996 Turnover
1995 Business Results	-.041	-.080	-.096	-.056
1996 Business Results	-.300	-.199	-.322*	-.272
1995 Customer and Market Focus	-.108	-.104	-.170	-.115
1996 Customer and Market Focus	-.174	-.159	-.169	-.154
1995 HR Development and Mgmt.	-.076	-.133	-.075	-.084
1996 HR Development and Mgmt.	-.112	-.048	-.101	-.087
1995 Information and Analysis	-.182	-.219	-.245	-.189
1996 Information and Analysis	-.227	-.220	-.265	-.225
1995 Leadership	-.162	-.217	-.242	-.185
1996 Leadership	-.276	-.253	-.314*	-.264
1995 Process Management	-.099	-.177	-.103	-.106
1996 Process Management	-.297	-.252	-.367*	-.277
1995 Strategic Planning	-.099	-.147	-.207	-.111
1996 Strategic Planning	-.097	-.111	-.115	-.094

Note. N = 29 or 32.

* $p < .10$

Table 2

Correlations between A Priori Scales and Grievances and Turnovers for 1995 and 1996

A Priori Scale	1995 Grievances	1996 Grievances	1995 Turnover	1996 Turnover
1995 Attitudes/Morale	-.128	-.203	-.226	-.115
1996 Attitudes/Morale	-.298	-.197	-.319*	-.261
1995 Awareness of Strategic Challenge	-.099	-.147	-.207	-.111
1996 Awareness of Strategic Challenge	-.097	-.111	-.115	-.094
1995 Communications	-.299*	-.363**	-.310*	-.303*
1996 Communications	-.298	-.244	-.325*	-.291
1995 Leader's Visible Commitment to Goals	-.192	-.245	-.251	-.205
1996 Leader's Visible Commitment to Goals	-.325*	-.303	-.372**	-.313*
1995 Supervisor's Concern for Improvement	-.068	-.137	-.031	-.066
1996 Supervisor's Concern for Improvement	-.292	-.241	-.386*	-.275
1995 Cooperation	-.106	-.160	-.189	-.132
1996 Cooperation	-.223	-.216	-.245	-.211
1995 Customer Orientation	-.158	-.167	-.225	-.170
1996 Customer Orientation	-.174	-.159	-.169	-.154
1995 Innovation	-.010	-.082	-.141	-.038
1996 Innovation	-.240	-.184	-.321*	-.225
1995 Involvement	-.022	-.038	-.038	-.025
1996 Involvement	-.282	-.227	-.323*	-.269
1995 Leader's Involvement	-.194	-.203	-.246	-.205
1996 Leader's Involvement	-.200	-.224	-.181	-.176
1995 Quality Policy/Philosophy	-.019	-.051	-.134	-.052
1996 Quality Policy/Philosophy	-.122	-.087	-.171	-.126
1995 Awareness of Productivity/Quality Issues	-.009	-.037	-.049	-.001
1996 Awareness of Productivity/Quality Issues	-.256	-.175	-.259	-.216

(table continues)

Table 2 (continued)

A Priori Scale	1995 Grievances	1996 Grievances	1995 Turnover	1996 Turnover
1995 Supervisor's Role in Quality Improvement	-.120	-.200	-.169	-.138
1996 Supervisor's Role in Quality Improvement	-.258	-.259	-.320*	-.250
1995 Rewards/Recognition	-.076	-.133	-.075	-.084
1996 Rewards/Recognition	-.112	-.048	-.101	-.087
1995 Social Interactions	-.199	-.248	-.175	-.207
1996 Social Interactions	-.276	-.135	-.343*	-.269
1995 System/Structure for Quality Improvement	-.001	-.002	-.030	-.001
1996 System/Structure for Quality Improvement	-.356*	-.346*	-.386**	-.320*
1995 Task Characteristics	-.166	-.175	-.229	-.166
1996 Task Characteristics	-.238	-.257	-.273	-.244
1995 Value Systems/Ethics	-.171	-.253	-.249	-.219
1996 Value Systems/Ethics	-.274	-.286	-.301	-.271
1995 Vision for the Future	-.144	-.186	-.197	-.154
1996 Vision for the Future	-.244	-.187	-.290	-.241
1995 Perceptions of Work Environment	-.063	-.089	-.119	-.087
1996 Perceptions of Work Environment	-.177	-.103	-.131	-.096

* $p < .10$

** $p < .05$

A CLEARANCE STUDY OF NITROTYROSINE FROM A PROSTATE CANCER CELL LINE

Vanessa D. Le
Graduate Student
Division of Pharmaceutics

University of Texas at Austin
College of Pharmacy
Austin, Texas 78712-1074

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1997

A CLEARANCE STUDY OF NITROTYROSINE
FROM A PROSTATE CANCER CELL LINE

Vanessa D. Le
Graduate Student
Division of Pharmaceutics
College of Pharmacy
University of Texas at Austin

Abstract

Formed by the reaction of superoxide with nitric oxide, peroxynitrite has been implicated in interfering with normal signal transduction pathways via its formation of nitrotyrosine. In carrying out the following experiment, our aim was to determine the clearance of nitrotyrosine from a human metastatic prostate cancer cell line, specifically, LnCap cells. In particular, we hope to deduce any specific mechanisms that may be involved in the clearing of nitrotyrosine which may be indicative of peroxynitrite repair. To this end, we needed to establish the smallest possible dose of PN that is required, but yet, not be lethal to the cell. The possible dose was determined from a previous pharmacodynamics study that was conducted of peroxynitrite in LnCap cells, here, we were able to establish that there is a significant (< 0.05 relative to control) loss of membrane integrity only at the 1mM PN dose 1 hr (7.7%) and 4 hour (23%) after exposure but not at 24h. Thus, following a 1mM PN exposure, cell numbers and viability were determined over a period of 3 days and quantitation of the extent of protein nitration was determined via a flow-cytometry protocol.

A CLEARANCE STUDY OF NITROTYROSINE FROM A PROSTATE CANCER CELL LINE

Vanessa D. Le

Introduction

The rapid reaction of nitric oxide with superoxide anion yields the powerful, cytotoxic oxidant peroxynitrite, which reacts with a variety of biomolecules in vivo and in vitro.[1,2,3] At physiological pH, PN has a half life of about 1 second in phosphate buffer and takes on a protonated form, whereby, it rapidly decomposes to form highly reactive species known to initiate lipid peroxidation as well as hydroxylate and nitrate aromatic residues.[4,5] Specifically, the toxicity of peroxynitrite has been attributed to its ability to nitrate tyrosine residues on nearby proteins and targets them for proteolytic degradation,[6,7] thereby, interfering with signal transduction by trophic factors, which are in part mediated by tyrosine kinases.[8] Due to the increasingly important role of peroxynitrite on the inflammatory response, it is our primary interest to study the clearance mechanisms of this cytotoxic oxidant from a cellular model, which underlies the focus of this paper.

Materials and Methods

Synthesis of peroxynitrite. Peroxynitrite was synthesized as previously described.[9] Briefly, an ice-cold 1M $NaNO_2$ solution plus a 1 M H_2O_2 solution is rapidly stirred in a 250-mL beaker. A 0.3M HCl solution is rapidly added to the nitrite/peroxide solution followed by the addition of 1.4M NaOH approximately 1 sec later. Unreacted hydrogen peroxide was removed by addition of MnO_2 to the solution. The concentration of peroxynitrite was spectrophotometrically determined by measuring the absorbance at 302 nm [E_{302} = 1,670 M^{-1} cm^{-1}].[10]

Cell culture. For nitrotyrosine and trypan blue exclusion experiments, following PN exposure, LnCap cells were seeded at a density of 1 x 10^6 cells on 90-mm petri dishes in RPMI 1640 medium supplemented with 10% fetal bovine serum, 50 units/mL of penicillin, 25 ug/mL of streptomycin, and 0.3 mg glutamine per mL. All chemical reagents and cell culture supplies were obtained from Sigma Chemical Co. unless otherwise noted. For MTT and SRB experiments, cells were seeded at a density of 5.82×10^3 cells per well in 96-well plates following PN

exposure.

Treatment with peroxynitrite. LnCap cells were harvested from 3 T-75 flasks (with approximately 90% confluency). Following trypsinization, cells were placed in each of two 50-mL centrifuge tubes and resuspended in 5-mL of Solution A buffer (5mM KCl, 50 mM Na_2HPO_4 1 mM $CaCl_2$, 0.8 mM $MgCl_2$, 90 mM NaCl, and 5 mM glucose, pH 7.4) for 5 minutes before addition of a single bolus of peroxynitrite. PN dosing solution was prepared using a 1:100 dilution of stock PN solution and 10 mM NaOH. The PN dosing solution was rapidly added to the cell suspension to achieve a final concentration of 1 mM. Cells were then spun down and resuspended in 3 mL of RPMI 1640 medium and seeded onto petri dishes and 96-well plates at appropriate density.

Measurement of mitochondrial respiration. Mitochondrial respiration was determined by the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) assay from Sigma Chemical Co.[11] Briefly, cells seeded at a density of 5.82×10^3 in a 96-well plate were incubated for two hours in medium containing 0.42 mg/mL MTT, thereafter, the medium was removed and 1 mL of dimethyl sulfoxide was then added. Within 30 min of DMSO addition, the absorbance was read at 492 nm on a 96-well microtiter plate reader following a brief 30-sec shaking time.

Sulforhodamine B (SRB) Assay. At various time following PN exposure, cells seeded at a density of 5.82×10^3 in a 96-wells plate were stained with sulforhodamine B and analyzed spectrophotometrically for total protein. The SRB assay measure the amount of total protein adherent to the bottom of the well after fixation with trichloroacetic acid. This method serves as an indirect measure of cell number, whereby, cells killed by exposure to changes in atmospheric pressure may adhere to the bottom of the well for some time before detaching.[12] Briefly, 50uL of cold 50% w/v TCA was added to each well and incubated for 1h at 4°C. The cells were washed 5 times with deionized water and stained for 10 minutes at room temperature with 25uL/well SRB, 0.4% w/v in 1% acetic acid. Plates were washed 5 times with 1% acetic acid and left to air dry overnight. The following day, protein-bound dye was solubilized with 200 uL of 10mM Tris-base, pH 10 and the absorbance was read with a microtiter plate reader at 492 nm.

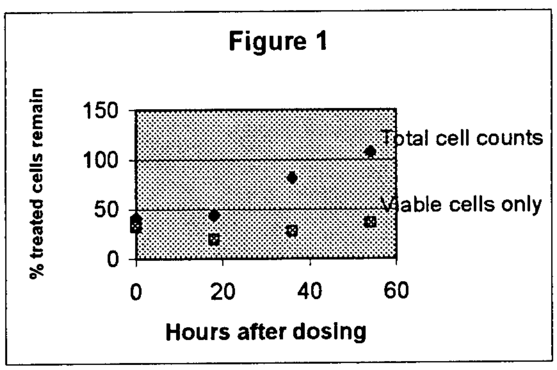
Trypan blue exclusion assay. Using a Zeiss Axiovert phase-contrast microscope, cell numbers were determined at various time following PN exposure using a 50:25 dilution of cell solution with trypan blue dye. Viable cells will

maintain their usual color, whereas, nonviable cells will incoporate trypan blue as a result of membrane damage and appears blue under the microscope.

Determination of cell morphology. At various times following PN exposure, cells were harvested and stained with propidium iodide to determine the fraction of cells that were apoptotic.

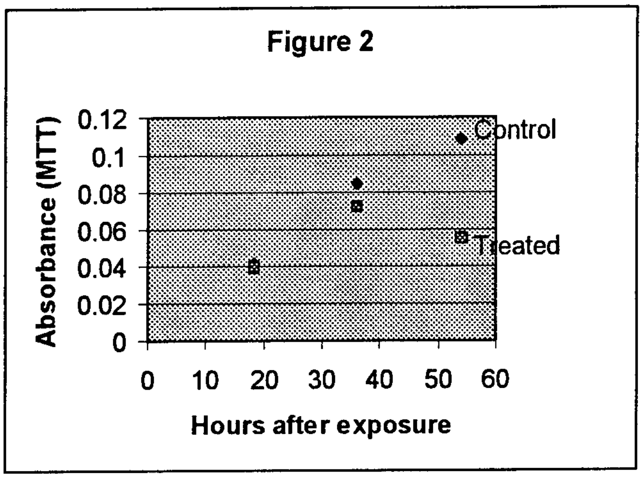
Mesurement of protein nitration. At various times following PN exposure, cells were harvested and fixed in ice cold 70% ethanol for 15 min. After decanting the supernatant, cells were permeabilized in 1 mL of ice-cold 0.25% NP-40. Cells were then spun and resuspended in 10% FBS, which blocks for non-specific protein binding. Cells were then exposed to an anti-nitrotyrosine rabbit ploycloncal antibodly (Upstate Biosciences) 1:20 dilution with 1% BSA for 1.5h at room temperature. Cells were washed twice with 1 mL of 10% FBS and exposed to goat anti-rabbit FITC conjugated antibody (Sigma) 1:20 dilution with 1% BSA for 30 min at room temperature. Cells were then run on a FACScan flow cytometer and 15,000 SSCxFSC gated events were evaluated from FL-1. Gates used for exclusion of debris were established in control sample for each experiment and applied to all treated samples. An increase in mean FL1 signal corresponds to an increase in the amount of nitrotyrosine. [13]

Results

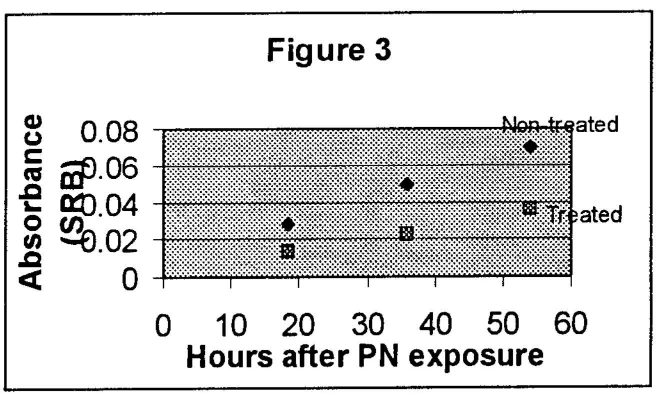


Referring to Figure 1, immediately after LnCap cell cultures were treated with a single bolus of 1 mM peroxynitrite, the percentage of total cells remain, as determined by the trypan blue exclusion assay, was 52% and 41% for the non-treated and treated samples, respectively. When considering only viable cells, the percentage declined to 36% and 33% for non-treated and treated samples, respectively. 18 hours after dosing, non-treated samples saw a significant increase in cell counts to 87%, whereas, the treated samples saw only a slight increment

to 44%. The percentage of viable cells was 41% and 20% for non-treated and treated samples, respectively. 36 hours after exposure saw a significant increase in total cell numbers for treated samples to 82%; whereas, the percent of viable cells account for only 28% of the treated samples. 54 hours after dosing, the percent of total cell counts has increased to 108%, yet, the percent of viable cell counts accounted for only 36%.



In reference to Figure 2, untreated samples showed a gradual increase in cell counts with increasing incubating times as expected. However, in the treated samples, there appears to be pronounced cell deterioration after an incubation period of 36 hours. This could be attributed to the effects of apoptosis on these cells, whereby, a delayed cell death is observed after 36 hours of PN treatment.



The data presented in Figure 3 indicates that at about 18 hours after PN treatment, the treated cells was able to overcome the acute toxic effects of PN and proliferate at a constant rate, similar to the growth pattern of non-treated cells. As presumed, there appears to be more activities in the non-treated samples as compared to the treated samples, thus, this is indicative that the cell counts are higher for non-treated samples than in the treated samples. An important point to note here is that although both the treated and the non-treated cells are growing at a constant steady rate, the non-treated cells are evidently proliferating at a higher rate than does the treated cells.

The following table summarizes the results obtained from the flow cytometry for the nitrotyrosine assay.

Table 1: Mean value of channel in FL1 corresponding to FITC. Mean of gated signals taken from FL1 x FSC plot. CV is in parenthesis.

Samples	Hours after PN treatment	Mean (CV)
Control, no primary Ab	0	10 (44)
Control, with primary Ab	0	21 (270)
Treated, no primary Ab	0	11 (45)
Treated, with primary Ab	0	347 (155)
Control, no primary Ab	18	154 (49)
Control, with primary Ab	18	315 (67)
Treated, no primary Ab	18	37 (53)
Treated, with primary Ab	18	381 (110)
Treated, no primary Ab	36	21 (46)
Treated, with primary Ab	36	1294 (165)
Treated, no primary Ab	54	26 (38)
Treated, with primary Ab	54	1311 (177)

As mentioned previously, an increase in mean FL1 signal corresponds to an increase in the amount of nitrotyrosine. The data presented in Table 1 shows that with increasing incubating time after PN treatment, the treated samples labeled with primary Ab saw a profound increase in nitrotyrosine signals; this is particularly evident between the 18 and 36 hour samples where there was a 9 fold difference. The data also indicates that the addition of primary Ab to the treated samples saw a significant increase in nitrotyrosine signals over the treated samples without primary Ab addition. There appears very little difference in signals between the 36 and 54 hour

incubation period. This discrepancy could be attributed to the fact that by 36 hours after PN treatment, LnCap cells were able to overcome the intial PN insult and were able to divide, hence, it then follows that the level of nitrotyrosine would decline as a result of this.

Discussion

Numerous published reports have provided strong indications that nitrotyrosine is associated with the induction of oxidative stress in human and animal models of disease.[14,15,16,17,18,19,20] In a recent report, Gow and co-workers provide strong indications that there exist some proteolytic and degradative pathways to account for the loss of nitrotyrosine in fixed and freshly prepared tissue samples.[21] With this in mind, our hypothesis that there exists some specific mechanisms for the clearance of nitrotyrosine from LnCap cells was tested using 1) MTT, SRB and trypan blue exclusion assays to account for cell viability after peroxynitrite treatment and 2) a flow cytometric assay to measure the extent of nitrotyrosine formation.

Our results shows strong evidence that there indeed exists an explanation for the clearance of nitrotyrosine from a cellular model, specifically, LnCap cells. The viability assays shows a profound cell loss immediately after PN treatment (59%), yet, by 18 hour after exposure to PN, the cells were able to overcome any toxic insult imposed by peroxynitrite and are able to divide and proliferate. It logically follows that as the number of cells increases, the nitrotyrosine signals decline as time progresses. Hence, by 54 hours, full recovery is in progress and the cells exhibit a relatively weak signal. The percent cell loss immediately after treatment reported here is slightly different from what was reported in past literature, such that, a concentration of approximately 1mM PN results in 50% cell death.[8] The discrepancy here could be attributed to the different cell line under study and also treatment conditions may have varied.

An important consideration to bear in mind when working with such an unstable reagent, such as peroxynitrite, is that PN decomposes with a half-life of about 1 second at physiological pH. Consequently, much peroxynitrite decomposes during mixing with the buffer before it ever contacts the cells. This important factor contributes to much variation amongst experiments. The data reported here is rather inconclusive at this point;

more experiments need to be conducted to reaffirm the findings as well as to test whether or not the data are reproducible. As it stands now, the data presented here was from a single experiment, thus, statistical analysis could not be performed.

References

1. Beckman, J.S., Beckman, T.W., Chen, J., Marshall, P.A. and Freeman, B.A. (1990) Proc. Natl. Acad. Sci. USA 87, 1620-1624.
2. Koppenol, W.H., Moreno, J.J., Pryor, W.A., Ischiropoulos, H. and Beckman, J.S. (1992) Chem. Res. Toxicol. 5, 834-842.
3. Pryor, W.A. and Squadrito, G.L. (1995) Am. J. Physiol. (Lung Cell. Mol. Physiol. 12) 268, L699-L722.
4. Radi, R., Beckman, J.S., Bush, K.M. and Freeman, B.A. (1991) J. Biol. Chem. 266, 4244-4250.
5. Radi, R., Beckman, J.S., Bush, K.M. and Freeman, B.A. (1991) Arch. Biochem. Biophys. 233, 481-487.
6. Ischiropoulos, H., Zhu, L., Chen, J., Tsai, J.-H.M., Martin, J.C., Smith, C.D. and Beckman, J.S. (1992) Arch. Biochem. Biophys. 298, 431-437.
7. Ischiropoulos, H. and Al-Mehdi, A.B. (1995) FEBS Lett. 364, 279-282.
8. Estevez, A.G., Radi, R., Barbeito, L., Shin, J.T., Thompson, J.A. and Beckman, J.S. (1995) J. Neurochem. 65:4, 1543-1550.
9. Beckman, J.S., Chen, J., Ischiropoulos, H. and Crow, J.P. (1994a) Methods Enzymol. 233, 229-240.
10. Hughes, M.N. and Nicklin, H.G. (1968) J. Chem. Soc. [A] 1968, 450-452.
11. Gross, S.S. and Levi, R. (1992) J. Biol. Chem. 267, 25722-25729.
12. Skehan, P., Storeng, R., Scudiero, D., Monks, A., McMahon, J., Vistica, D., Warren, J.T., Bokesch, H., Kenney, S. and Boyd, M.R. (1990) J. Natl. Cancer Inst. 82:13, 1107-1112.
13. Kalns, J.E. , Parker, J., Holwitt, E., and Kiel, J., in press.
14. Salman-Tabchem, S., Guerin, M.C. and Torreilles, J. (1995) Free Rad. Biol. Med. 19, 695-698.
15. Schulz, J.B., Matthews, R.T., Jenkins, B.G., Ferrante, R.J., Siwek, D., Henshaw, D.R., Cipollini, P.B., Mecocci, P., Kowall, N.W., Rosen, B.R. and Beal, M.F. (1995) J. Neuroscience 15, 8419-8429.
16. Szabo, C., Salzman, Al.L. and Ischiropoulos, H. (1995) FEBS Lett. 363, 235-238.
17. Ischiropoulos, H., Al-Mehdi, A.B. and Fisher, A.B. (1995), Am. J. Physiol. 269, L158-L164.
18. Haddad, I.Y., Pataki, G., Hu, P., Beckman, J.S. and Matalon, S. (1994) J. Clin. Invest. 94, 2407-2413.
19. Beckman, J.S., Ye, Y.Z., Anderson, P.G., Chen, J., Accavitti, M.A., Tarpey, M.M. and White, C.R. (1994) Biol. Chem. Hoppe-Seyler 375, 81-88.
20. Gow, A.J., Duran, D., Malcolm, S. and Ischiropoulos, H. (1996) FEBS Lett. 385, 63-66.

THE EFFECT OF 2.06 GHz MICROWAVE IRRADIATION ON THE PERMEABILITY OF THE BLOOD BRAIN BARRIER

Kavita Mahajan
Graduate Student

Trinity University
715 Stadium Drive
San Antonio, TX 78212

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office Of Scientific Research
Brooks Air Force Base, TX

and

Armstrong Laboratory

August 1997

THE EFFECT OF 2.06 GHz MICROWAVE IRRADIATION ON THE PERMEABILITY OF THE BLOOD BRAIN BARRIER

Kavita Mahajan
Graduate Student
Trinity University

ABSTRACT

It has been demonstrated that the intact nature of the Blood Brain Barrier (BBB) has the ability to be compromised by exposure to electromagnetic field radiation. Previous studies were conducted in an effort to determine whether the induced breakdown of the BBB may be attributed to microwave irradiation or hyperthermia. In our study, we hypothesized that the permeability of the BBB may be altered by an induced hyperthermia as the result of exposure to microwave irradiation. Four groups of rats were created based on the method of Albumin permeability analysis. Groups included qualitative analysis of Sodium Flourescein (4%) and Evan's Blue, immunocytochemical analysis of Albumin and spectrophotometrical analysis of 0.5% Sodium Flourescein. Rats were exposed to 2.06 GHz microwave irradiation at power densities of 40 mW/cm^2 and 130 mW/cm^2. In addition to experimental animals, positive control and sham- irradiated animals were included in each group. It was found that power densities less than 130 mW/cm^2 were unable to induce a hyperthermic event and no Evan's Blue tracer was detectable in the brain. However, exposure to 130 mW/cm^2 induced leakage of Evan's Blue and was additionally accompanied by tympanic temperatures above 43.2°C. Presently, results are inconclusive based on the incomplete studies of the Sodium Flourescein and Albumin subjects.

THE EFFECT OF 2.06 GHz MICROWAVE IRRADIATION ON THE PERMEABILITY OF THE BLOOD BRAIN BARRIER

Kavita Mahajan

INTRODUCTION

The blood brain barrier (BBB), the brain's natural barrier system, is composed of the continuous membranes of the endothelial cells comprising the capillary beds of the brain. The tight junctions existing between adjacent cells create a selectively permeable membrane between the blood and the interstitial spaces of the brain, thus regulating the permeability of plasma borne solutes and cerebrospinal fluid elements into the brain's extracellular fluid (1). This highly regulatory nature of the barrier functions to maintain tissue stability of cations, enzymes, proteins and lipids within the brain, allowing only a controlled selection of substances to enter the organ. Great fluctuations in the concentration of blood solutes requires a strictly controlled system to sustain the equilibrium within the brain (2). The integrity of the BBB is therefore essential in providing the appropriate conditions for strict control of the brain's fluid environment.

Numerous conditions, both natural and induced, exist to alter the permeability of the BBB (3). The degree of permeability due to this alteration has been shown to be directly related to the extent of breakdown of the intact BBB. Specifically, in addition to conditions such as hypertonic solutions, metal poisoning, and auto-immune diseases, it is speculated that changes in the BBB may be induced by exposure to low level electromagnetic field (emf) radiation (1). The increased permeability of the usually highly selective membrane is hypothesized to occur via a mechanism of hyperthermia, whereby elevated brain temperatures are induced by exposure to microwave radiation. Several studies have been conducted in an effort to demonstrate this effect. However, as of present, it has not been possible to successfully duplicate experiments which support this model of permeability change due to hyperthermia.

A study conducted by Frey et al (1975) demonstrated the permeability of Sodium Fluorescein (Na

Fluorescein) through the blood brain barrier subsequent to 1.2 GHz radiation exposure at power densities of 0.2 mW/cm^2 and 2.4 mW/cm^2. The presence of fluorescence within these experiments was not reported to be accompanied by any hyperthermic event (4). Additionally, Merritt et al reported that fluorescence of the brain parenchyma was not visible or could not be attributed to microwave induced (1.2 GHz: 2 mW/cm^2- 75 mW/cm^2) penetration of Na Fluorescein into extracellular fluid (1). However, data from this study included positive results from a hyperthermic model whereby subjects were exposed to environmental heat conditions until brain and colonic temperatures increased to the same level as would have been observed via heating with 75 mW/cm^2 microwave radiation. Increased BBB permeability was thus concluded to be an effect of hyperthermia only and not microwave exposure.

Furthermore, a study conducted by Lin and Lin (1980) support this model of hyperthermic induced breakdown of the BBB. Exposure to 2450 MHz at power densities from 0.5 to 2600 mW/cm^2 failed to induce breakdown of the blood brain barrier. An increase in the power density to 3000 mW/cm^2 was accompanied by an increase in the amount of Evan's blue tracer present in the brain tissue. Additionally, it was also noted that this event involved an increase in brain temperature to an excess of 43°C (5). Replication of this experiment using Na Fluorescein showed similar results: no tracer was present in the brain at cortical temperatures below 41°C (6). Thus, in conjunction with Merritt's results, it is hypothesized that increased permeability of the BBB may be attributed *only* to hyperthermia induced within the brain parenchyma due to heating, but not specifically to microwave irradiation.

METHODS

Exposures:

All irradiated rats were exposed to 2.06 GHz radiation in K polarization (head facing towards horn) using an L-Band Klyston source (Model 2852, Colber Electronics, Stanford, CT) in an anechoic chamber (Emerson and Cuming, Inc., Canton, MA). Irradiation was conducted in the far field (1.65 m) at power densities of 40 mW/cm^2 and 130 mW/cm^2. Exposure orientation and distance were kept constant

by placing rats on a styrofoam table. Additionally, sham exposed rats were monitored in the chamber, while positive control rats were monitored in the surgery area. Colonic and right tympanic measurements were recorded for irradiated rats and colonic measurements only for sham- irradiated rats using non-metallic rectal probes (+/- 0.1° C, Model 101, Vitek, Boulder, CO).

Infusions:

Rats underwent 1 of 4 treatments prior to sham exposure or irradiation. Sodium pentobarbital (50 mg/ kg, IP) was administered prior to surgery. A catheter was inserted into the right femoral vein for infusion of treatments.

The Evan's Blue protocol (n=4) included infusion of 2% Evan's Blue (0.1 ml/ 100 gm) into the femoral catheter over period of one minute, followed by a 0.2 mL chase with 0.9% saline. Temperature measurements were collected during exposure or sham irradiation which was conducted for a period of 30 minutes followed by a 20 minute equilibration period. Subsequent to exposure, rats were removed from the chamber and perfused intracardially using 0.9% NaCl and 4% paraformaldehyde/ 0.1% Phosphate buffer saline/ 0.01% thimerosal. Whole brains were removed, stored in sucrose, sliced (40 um sections) and viewed with fluorescent microscopy.

The Sodium Fluorescein protocol (n=4) was identical to the Evan's Blue procedure except that 4% Sodium Fluorescein (1 cc) was substituted for the Evan's Blue infusion. Exposure and brain slicing parameters were identical.

A quantitative analysis of Sodium Fluorescein tissue concentrations was conducted via collection of *tissue punch* samples. Subjects were infused via the femoral vein with 1 cc of 0.5% Sodium Flourescein and exposed identically to other subjects. The animals were perfused and selected brain regions were isolated. Samples were weighed and homogenized in 300 ul butanol and centrifuged at 1500 rpm for 5 minutes. The supernatant was removed and placed in a tube containing 1800 ul of borate buffer

(pH= 10). Samples were then vortexed and centrifuged again. Subsequent to the second centrifugation, the lower buffer phase was isolated and analyzed via spectrophotometry at 482 nm excitation and 518 nm emission.

A study of Albumin penetration of the blood brain barrier was carried out via immunocytochemical analysis. A mock surgery was conducted on exposure and non-exposure subjects. The rats were intracardially perfused with saline and paraformaldehyde, and brains stored in paraformaldehyde for 12 hours followed by transfer to sucrose until the time of slicing. 40 um sections were obtained and stored in 0.01 M Phosphate Buffered Saline (PBS) until immunocytochemistry was performed.

In addition to exposed and non-exposed groups, a positive control group was included with each treatment. Preparation included insertion of a femoral vein catheter (with constriction of the right external carotid artery) and infusion of 1 cc Evan's Blue or 1 cc Sodium Flourescein. After five minutes, 10 M urea (0.1 ml/100 g body weight) was infused into the right internal carotid artery and the animal was monitored for a 20 minute equilibrium period until perfusion. Albumin positive control subjects received only an infusion of urea into the right internal carotid artery.

Immunocytochemistry

Albumin analysis was conducted via staining with albumin antibody. Sections were incubated in a solution of 1.0% horse serum/ 0.01M PBS/ 0.2% Triton X-100 for 20 minutes followed by rinse with 0.01 M PBS. Incubation in albumin primary antibody (rabbit anti-rat IgG, Organon Teknika, Durham, NC) followed for 24 hours @ 4° C. Sections were rinsed in 0.01 M PBS and incubated in biotinylated secondary antibody (Rabbit IgG, PK 4001, Vector Laboratories) for 60 minutes at room temperature on a shaker. Incubation in Vectastain ABC mixture followed for 60 minutes at room temperature. Sections were rinsed in 0.01 M PBS, and placed in 0.03% 3,3'diaminobenzidine tetrahydrocholoride (Sigma)/ 0.0015% Hydrogen peroxide for 1-3 minutes or until sections were evenly stained. Immediate rinsing

with distilled water followed. Sections were mounted on gelatin coated slides and coverslipped.

Brain Temperature Measurements

Rats were anesthetized with Ketamine (70mg/ kg IP) and Xylazine (10mg/kg IP) and stereotaxic surgery was performed. Coordinates for insertion of Vialon guides (Becton Dickinson) are as follows: 1.8 posterior to bregma, 1.5mm lateral from midline, and 8.3 mm below dura. Guides were secured by cranioplastic cement (Plastic One, Roanoke, VA) and sealed to nylon screws. Thermal probes (Vitek, Boulder, CO) were inserted through these guides to record hypothalamic and cortical temperatures. Methods of this procedure have been described previously in detail (7,8).

RESULTS

Presently, results of this study are inconclusive due to difficulties with immunocytochemical analysis of albumin penetration in the brain and the continued ongoing study of Sodium Fluorescein treated subjects (both microscopy and spectroscopy groups). The only groups available for assessment at the current time are the Evan's Blue treated subjects. Colonic and tympanic temperatures were measured for each exposure or non-exposure group (n=4).

Fluorescent microscopy analysis of Evan's Blue brain sections yielded qualitative results that correlate with colonic and tympanic temperatures. Positive control rats (i.e. treated with urea) demonstrated significant extravasation of Evan's Blue into the right hemisphere of the brain. Fluorescence was especially marked in the cerebral cortex, caudate, and hippocampus. Additionally, a greater intensity of fluorescence was noted in the anterior regions of the brain in comparison to posterior regions.

Sham exposed rats showed only a minimal increase in colonic temperature and qualitative examination of brain sections yielded negative results of Evan's Blue fluorescence. Subjects exposed to 40 mW/cm^2 demonstrated only a slight increase in tympanic and colonic temperature. Microscopy

indicated that there was no penetration of Evan's blue into either hemisphere of the brain with 40 mW/ cm^2 exposure. Qualitative review of brain sections of subjects exposed to 130 mW/cm^2 yielded positive but faint results in both hemispheres of the forebrain which was accompanied by a significant increase in colonic and tympanic temperatures (Figs 1 and 2).

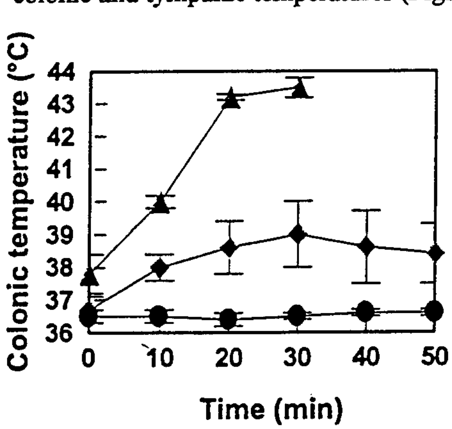


Fig. 1

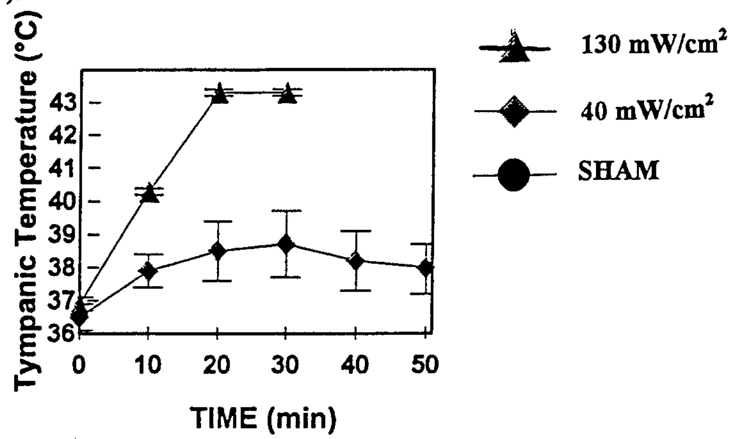


Fig. 2

Brain temperatures upon exposure of non-subject rats (no treatment with Evan's Blue) to 0.8 mW/cm^2 were tracked via hypothalamic temperature measurements. Mean hypothalamic temperature (T_{hyp}) at baseline was recorded at 35° C. Mean T_{hyp} after 30 minutes exposure to 0.8 mW/cm^2 was recorded at 34.8° C, thus indicating a negative rate of rise. Additionally, non- subject rats exposed to 40 mW/cm^2 were analyzed for changes in hypothalamic and cortical temperatures. Results showed that mean tympanic temperature (T_{tymp}) at baseline were measured at 36.38 °C. After 30 minutes of heating mean T_{tymp} was measured at 39.18 °C. Mean T_{hyp} at baseline was 35.25 °C, and T_{hyp} at 30 minutes was measured at 38.15 °C.

Exposure to 130 mW/cm^2 induced increased hypothalamic temperature to 41°C upon which exposure was discontinued. Mean T_{tymp} at baseline was measured at 37.1° C and temperature after T_{hyp} reached 41° C was 40.9°C. Mean T_{hyp} at baseline was 36° C.

CONCLUSION

Based on past experimental evidence, it was noted that Evan's Blue tracer could be detected in the brain only in subjects whose brain temperatures exceeded 43°C (5). Additionally, a separate experiment indicated penetration of Sodium Fluorescein into the brain only when cortical temperatures exceeded 41°C. (6). Temperature levels in both studies may be defined as the occurrence of a hyperthermic event.

As expected, results of this study demonstrate positive results of Evan's Blue extravasation into the brain accompanied by microwave (130 mW/cm^2) induced hyperthermia (as defined by T_{tymp} exceeding 43.2 °C). Sham irradiation and 40 mW/cm^2 exposed animals did not undergo an increase in either T_{col} or T_{tymp} to levels above 40°C and thus, no penetration of Evan's blue was noted in the subjects' brain tissue. Brain temperature measurements more accurately confirm findings of increased brain temperatures due to microwave exposure. Thus, the Evan's Blue subject data confirms the findings of Merritt et. al. The increase in brain temperature to a hyperthermic level, as induced by microwave heating, is sufficient to cause a breakdown of the intact and highly regulated blood brain barrier.

Previous speculation regarding "hot spots" within the brain may be accountable for the results of Evan's Blue subjects. This explanation involves the heating of regions of the brain induced by microwave exposure, thus causing a 'location- specific" breakdown of the intact blood brain barrier, and leakage of Evan's Blue tracer into that selectively heated region. This explanation does not account for, however, the even extravasation of dye into both cerebral hemispheres.

Currently, all results of this study are inconclusive. Further analysis and correlation between Evan's Blue, Na Fluorescein, and Albumin subjects are to be continued and examined in an ongoing study.

REFERENCES

1. Merritt, J.H., Chamness, A.F., Allen, S.J. Studies on Blood Brain Barrier Permeability After Microwave Radiation. *Rad and Environm. Biophys.* 15: 367-377 (1978).

2. Oldendorf, W.H. Blood Brain Barrier Permeability to Drugs. *A. Rev. Pharmac.* 14: 239-248, 1974.

3. Goldstein, M., Backstron, T., Ohi, Y., Frenkel, R. The effects of Ca^{2++} ions on the C^{14} catecholamine biosynethesis from C^{14}-tyrosine in slices from the striatum of rats. *Life Sci..* 9: 919- 924, 1970.

4. Frey, A.H., Feld, S.R., Frey, B. Neural function and behavior: Defining the relationship. *Ann. N.Y. Acad. Sci.* 247: 433-438, 1975.

5. Lin, J.C. Mechanisms of microwave- induced Blood- Brain Barrier Alterations. Arlington, VA: ONR Physiology Program, 1980.

6. Lin, J.C., Lin, M.F. Studies on microwave and blood brain barrier interaction. *Bioelectromagnetics.* 1: 313-323, 1980.

7. Mason, P.A., Walters, T.J., Doyle, J.M., Kane, J.L., Escarciga, R., Purdy, K.E. Heat Shock Protein and GFAP expression after microwave exposure or warm water immersion. FASEB J.9: A546, 1995.

8. Walters, T.J., Ryan, K.L., Belcher, J.C., Doyle, J.M., Tehrany, M.R., and Mason, P.A. Regional brain heating during microwave exposure (2.06 GHz), warm water immersion, environmental heating, and exercise. *J Therm. Biol.*, Submitted.